INVESTIGATION OF DISTORTION OF DIVERS' SPEECH USING POWER SPECTRAL ESTIMATES BASED ON THE FAST FOURIER TRANSFORM

by

William Howard Bond

United States
Naval Postgraduate School

THESIS

INVESTIGATION OF DISTORTION OF DIVERS'
SPEECH USING POWER SPECTRAL ESTIMATES
BASED ON THE FAST FOURIER TRANSFORM

by

William Howard Bond
and
James Michael Myatt

June 1969

Thesis by: William H. Bond and James M. Myatt entitled "Investigation of Distortion of Divers' Speech Using Power Spectral Estimates Based on the Fast Fourier Transform." June 1969.

ERRATA

Page	Line	Change	To
58	8	Fig. 23 a-c	Fig. 23
61	6	Fig. 20 (a-c)	Fig. 22 a-c
61	2nd paragraph, insert as first line -- "The transient spectrum with the oral cavity is shown in Fig. 24."		
61	3rd paragraph, insert as first line -- "Fig. 25 shows the transient spectrum at 80 ft. for the first 2048 points."		

Investigation of Distortion of Divers'
Speech Using Power Spectral Estimates
Based on the Fast Fourier Transform

by

William Howard Bond
Major, United States Marine Corps
B.S., University of Texas, 1957

and

James Michael Myatt
Captain, United States Marine Corps
B.S., Sam Houston State College, 1963

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 1969

ABSTRACT

The problem of distortion in underwater communications peculiar to free divers and techniques for analysis of speech wave forms are discussed. The Fast Fourier Transform algorithm, selected to analyze shifts in formant frequencies due to restricted oral cavities, high ambient pressures, and forced speech is discussed. The Fast Fourier Transform is used to analyze a vowel sound and show that the expected shifts do occur. Recommendations are made for extending the techniques to all non-noise like sounds and breathing mixtures other than compressed air.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The authors wish to thank the Communications Sciences Laboratory, Department of Speech, University of Florida for stimulating their interest in the thesis subject. Mr. Jim Lasch, Hydro-Products, Inc., Dillingham Corporation, San Diego, California, Major Alex Lee, U. S. Marine Corps, Marine Corps Development and Educational Center, Quantico, Virginia, and LCdr. Al Spinks, U. S. Navy, Navy Special Warfare Group, Coronado, California, assisted and many times initiated the long term loan of equipment necessary to perform the research. Mr. Lance Gardner, Marine Physical Laboratory, University of California at San Diego and divers of the Fifth Force Reconnaissance Company, Camp Pendleton, California spent long hours diving in cold dark, and deep water.

For the good advice (many times mistakenly unheeded), red tape cutting, and thoughtful guidance in hours of difficulty, we extend our heartfelt gratitude to Dr. Harold Titus, Professor, Electrical Engineering Department, Naval Postgraduate School.

I. INTRODUCTION

Man's future needs dictate the development of the vast areas of the continental shelf for food, raw materials, and living space. To explore, map, and colonize this unknown a myriad of new equipment must be developed for explorers of this "last frontier." Today man has the capability to communicate reliably over the vast reaches of space; but free divers (Divers with no attachment to a habitat or surface station — untethered) operating in close proximity cannot communicate reliably to each other. One UDT officer summed up the state of the art by saying, "reliable communications between free divers consists of tugs on buddy lines." Unsatisfactory undersea communications result from poorly designed oral cavities, high ambient pressures, poor electronics, and exotic gas breathing mixtures. Current equipment has been further complicated by designers unfamiliar with the environment and hazards associated with the free diving world; few designers have taken into consideration the user's emotional response to the apparatus in the underwater environment.

In 1959-60, the problem of inadequate communications between free divers was identified in a specific operational requirement of Navy UDT and Marine Corps Force Reconnaissance Divers. Since that time, the U. S. Navy and U. S. Marine Corps have been moving on separate paths to develop a reliable communicator for the free diver. Presently available equipment is far from an optimum design for many reasons — poor overall intelligibility being of prime importance. In short, the state of the art is unacceptable. This unacceptable condition, personal interest, and the requirement of the Armed Forces for more sophisticated devices generated the original topic of this research — an investigation into speech intelligibility of free diver underwater communicators. After visiting with persons engaged in Navy sponsored work in this realm, it became apparent that more was involved than just better electronics. Inherent in improving underwater communications devices is the need for better understanding of what happens to speech in the undersea environment. This fact made it necessary to pursue a working knowledge of the disciplines of linguistics and speech processing. Repeated conferences with the Communications Sciences Laboratory, University of Florida stimulated the curiosity and study of these disciplines.

Past speech processing has taken the form of qualitative analysis using analog equipment, such as spectographs, and intelligibility studies using human listeners; however, few quantitative studies have been undertaken to isolate the major distortion problems. Many comparative studies of hard line and free diver communication systems have been and are being carried out, primarily using human speakers and listeners to evaluate system intelligibility using controlled word lists [1]. The spectograph also gives a measure of quantitative analysis to highly trained speech analysts.

The use of frequency analysis (a common tool of the electrical engineer) for simple waveform analysis is a proven technique. In the past, this technique, transforming the signal under consideration to the frequency domain using the Fourier transform required too great an expenditure of time and human effort for speech waveforms. The advent of digital computers with vast memory capabilities, now makes this technique a feasible method for speech waveform analysis. Transformation of the speech waveforms is accomplished using the discrete

Fourier transform. This transformation is reliable if the signal is band limited and occurs for a finite period of time. The Fast Fourier Transform (FFT) algorithm provides a rapid method for calculating the discrete Fourier transform of such signals [2]. This algorithm has many applications to digital filtering and controlling real time processes in the frequency domain. The value is apparent when considering signals that require 8,000 samples to provide complete coverage of frequencies of interest. It provides a real time saving of 300 to 1 over previous algorithms; analysis that previously took hours now takes seconds.

This is the approach taken in this paper — investigation of distortion in small underwater communicators using the FFT to process diver's speech. The final value can only be imagined; the analysis can be applied to isolate distortion in helium speech, build better helium speech unscramblers and help to prevent loss of vital undersea explorers, such as Berry Cannon.

A. FREE DIVER COMMUNICATION

Whenever man goes underwater, he needs communication. Tasks underwater are made easier and safer if communications exist between divers and/or surface stations. The military diver uses communications to carry out complicated tasks, to penetrate enemy held areas, and to effect rendezvous with underwater vehicles. Reliable underwater communications do not exist for all cases of interest to the U. S. Navy and U. S. Marine Corps.

Communications between free divers is a relatively new need; but as man goes deeper for longer periods of time the requirement will increase in importance. Sonar has been used for communications between submersibles and/or surface vessels for many years; small sonar sets are now performing the same service between divers, submersibles and/or surface stations. Many problems exist in the diver's case that are not of consideration in submarines. The more important include the addition of restricted oral cavities to the vocal tract, increased ambient pressures, increased noise levels, and the uses of exotic breathing mixtures.

1. Current Equipment

Initially, efforts to provide the diver with communication involved amplifying the voice and projecting it into the underwater environment. Water noises near transmission frequencies, low power, and exclusion of microphones and earphones (for low cost) contributed to poor intelligibility and short range making these systems impractical for most military uses [3]. Electromagnetic and electric field transmission have also been tried with limited success at short ranges.

Sonar has provided the most successful method of communication to date. Using crystals possessing electrostrictive properties, a sound pressure or acoustic wave is generated in to the water and received by a similar transducer. Sets of this design operate around a carrier frequency of 8KHz-42KHz (the lower carrier producing greater range of transmission) and utilize AM-SSB/SC transmission to lower noise reception and increase power. The Hydro-Products 811 and Bendix POC-2 (see Figs. 1 and 2) underwater communicators are industry's latest efforts in the field. Characteristics of these sets are shown in Table I.

Changes occur in these sets when placed on the diver. The omnidirectional transducer becomes somewhat directional due to acoustical dampening by the lungs, tanks, and the neoprene wet suit worn in cold water. Two divers operating in close proximity can not communicate when

Hydro Products 811
Underwater Communicator

FIGURE 1

Bendix PQC-2
Underwater Communicator

FIGURE 2

TABLE I

SET	HP 811	POC-2
Carrier Frequency	8.0875 KHz * AM-SSB/SC	8.0875 KHz * AM-SSB/SC
Electric Power Into Water	.5 watt *	.6 watt *
Range	Nominal 800 meter *	500 meter minimum *
Modes	Voice	Voice, Tone, and Interrupted Tone
Transducer	Electrostrictive (barium titamate)	Electrostrictive (lead zirconate)
Pattern	Omnidirectional	Omnidirectional
Microphone	Oral Cavity Microphone	Bone conduction (two)
Receiver	Bone Conduction (one)	Bone Conduction (two)
Intelligibility	No information (under test at U. of Fla.)	No information (under test at U. of Fla.)
Bandwidth	300-3000 Hz * (2700 Hz)	300-3000 Hz * (2700 Hz)

* Manufacturer's Data

the lung cavity is in the line of transducers. This is a serious defect when one user is unaware of the relative location of the second user.

The communicator bandwidth operates on the telephone principle — 300-3000 Hertz bandwidth will provide high intelligibility. Studies indicate for shallow depths that this bandwidth will provide intelligibility of 95+% [1]; but this is not necessarily true for all conditions encountered by the diver.

Evidence points to bone conduction as the primary source of hearing underwater [4]. All currently available communicators utilize bone conduction reception; one of the tested communicators used bone conduction transmission in place of the normal microphone in the oral cavity. Some tests indicate a definite loss in intelligibility when using bone conduction transmission; but no quantitative results support these conclusions. The use of a single hose regulator with bone conduction transmission caused the signal to be obscured by bubble noise passing the conductors.

These problems are not unique to the tested communicators but a far more important set of problems are those common to all communication systems for free divers.

2. Problem Areas

The most complex problem is that of speech distortion caused by the addition of a small cavity over the mouth. To date little basic research has been attempted in an effort to design an optimal oral cavity. The design must balance two conflicting variables — small volume to prevent low resonances near formants or articulated sound and buildup of carbon dioxide, and large volume to prevent increased pressure from affecting the speaker in articulation. The Bioengionics' Nautilus oral cavity (see Figure 3) was used to aid speaker articulation in this research. This oral cavity has been used successfully to increase intelligibility of systems in other tests and provides an excellent trade off in cavity design and diver safety [5].

Increased operating depth results in increases of ambient pressures. This changes the characteristics of the resonant cavities and the density of the breathing gas. When a diver is operating at pressures of four atmospheres (100 feet) or greater, the voice attains a nasal quality and causes considerable loss of intelligibility [5].

When exotic breathing gases are substituted for the primarily nitrogen-oxygen gas, the voice attains a quality best described as the "Donald Duck" effect. The use of exotic gases causes formant shifts due to the increased sound velocities. Unfortunately the shifts are non-linear; this makes correction a complex problem. The amount of distortion due to these effects is under study by conventional techniques by the Communications Sciences Laboratory, University of Florida [1], and other agencies [6,7,8].

Physiological problems are of interest to ensure that human engineering of communicators are acceptable to the diver when operating in an alien environment and commensurate with desired safety. The diver feels an almost personal relationship with his equipment. The communicator must require little or no change in equipment for the diver to accept it unhesitatingly. A prime example is the addition of the oral cavity, a much needed item for proper articulation. The original design of the Bendix PDC-2 excluded use of a cavity for this reason.

Bioengionics Nautilus Oral Cavity

FIGURE 3

Design of workable underwater communicators is feasible; but a great deal of research is necessary before a single system will overcome the vast number of problems. The most critical item is development of a method to solve the deep depth and exotic gas speech distortion and then microminaturization of the "unscrambler" for inclusion in the free diver's equipment.

B. SPEECH PROCESSING

To the engineer engaged in speech research, "speech processing" implies synthesis and analysis, automatic recognition and speech compression. The linguist will speak of the acoustical model in speech production or perception and/or various articulation and intelligibility testing. The language of speech research is foreign to the beginner; therefore, Appendix A includes a glossary of speech and linguistic terms.

As previously mentioned, the requirement exists for a technique that will enable the speech analyst to accurately determine the effects of an alien environment, such as that encountered by the diver, on the human voice.

Some of the techniques used for speech analysis are summarized with a brief description of the new "wrinkle" in analysis following.

1. Analog Techniques

a. Recognition Analysis

The vowel sounds have been the subject of intense research. The most common result has been the classification of the vowels by their formants, primarily the relation of the formants to each other, and has been labeled "Visible Speech" [9]. The device most commonly used is the "acoustic spectogram" where a visual representation of the spoken message or phoneme is displayed in frequency versus time. The intensity, a third parameter, is displayed by the darkness of the plot. Thus, the transient spectrum can be viewed. A spectogram of the word "at" is shown in Fig. 4. The chief disadvantage of the spectogram is the lack of resolution.

b. Power Spectrum Measurements

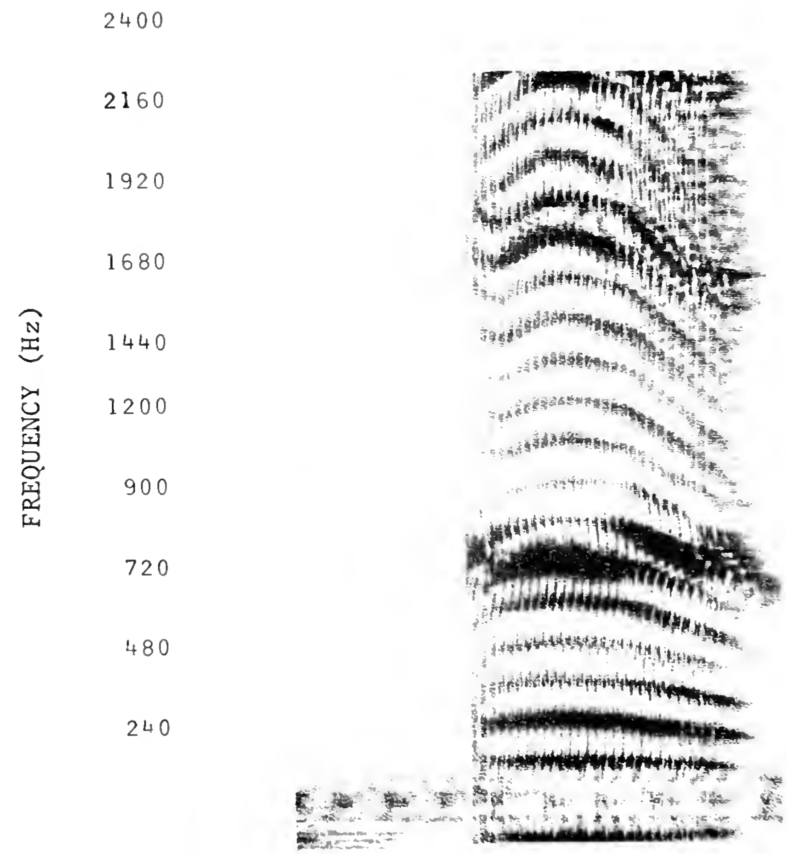
The analog spectrum analyzer is a convenient tool to view the spectrum of a signal instantaneously. It works on the principle of a constant width bandpass filter with a variable center frequency; however, high precision measurements are not possible. Another technique that borders on being classed as digital is the use of a bank of narrow bandpass filters with increasing bandwidth as their center frequency increases. Until recently, size and cost have limited their use, with resolution still the determining factor.

c. Waveform Analysis

The fact that vowel sounds can be characterized by their waveform asymmetry has been noted. Pairs of phonemes can be compared and separated on the basis of the magnitude and polarity of the output of a demodulator [10].

d. Articulation and Intelligibility Tests

To evaluate communication systems, the articulation tests were widely used during World War II. The results of the tests aided hardware design. Intelligibility testing takes the form of reading from phonetically balanced word lists and assigning a percentage score for the correct responses of human listeners. The consistency of such tests demonstrates their reliability and makes them the most commonly used; however, the search for optimum lists of sounds, words and phrases continues [5].



Spectogram of the Word "AT"

FIGURE 4

2. Digital Techniques

The analysis of speech by digital techniques owes a great deal of its popularity to the interest in the development of the "vocoder", or voice coder. Linguists and psycholinguists have been a trifle reluctant to master the techniques of the digital world [11].

Digital Spectrum analyzers prior to 1965 consisted of narrow band digital filters. The number of digital filters necessary for sufficient bandwidth and resolution required extensive memory and computing time; an excessive amount for speech waveforms.

a. The Fast Fourier Transform

The capability for high resolution spectral analysis was not available until recently discovered algorithms were implemented [12]. The algorithm of interest in this paper is the Fast Fourier Transform (FFT) [2]. It is an efficient algorithm for computing the "discrete Fourier transform". The number of operations is reduced from N^2 to $2N\log_2 N$, where N is the number of samples. Thus, it can be used for offline spectral analysis of processes. The advantage of the FFT is illustrated by comparing the number of operations required for N = 8192 samples. N^2 is in excess of 67 million compared to 213 thousand for the FFT. Such a saving in operations vastly reduces the computing time.

b. Spectral Measurements

With the FFT as a tool available in generating the Fourier coefficients, several avenues of analysis are plausible. The avenue chosen is the power spectrum. Other spectra, such as the cesptrum, are available but were not considered. The FFT lends itself to such uses and could be applied in future research in this area. By claiming that the power spectrum is available, it must be clear that the spectrum is only an estimate. How good the estimate is depends upon the methods of implementing the algorithm. First, the signal is of finite length, hence the number of samples is finite. Second, the signal is bandlimited. The assumption is made that the signal is a member of a stationary process.

II. DATA COLLECTION

After consultation with authorities in speech processing and underwater communications, and after considering time, funds, and human effort available, the decision was made to proceed to collect data to support two independent projects. This decision was influenced by the loan, free of cost, of the two underwater communicators used in the data collection. Certain control sounds and words, shown in Table II, were recorded to support the digital speech processing. Two vowel and four consonants were arbitrarily chosen. One sound of each main subgroup was included [9]. This data was originally intended to assist in isolating distortion in the two communicators. The second project was designed to assist in judging the relative intelligibility of the two systems under various conditions. Two Campbell word lists were recorded by each diver. The Campbell word lists are monosyballic, phonetically balanced lists of 25 words each, selected randomly from a master list of 500 words. Table III is an example of such a list.

Original plans called for data to be taken in two locations, an anechoic pool free of outside interference and the ocean to provide a noisy environment with varying conditions of range and depth. Sea conditions and boat availability in January and February, 1969 allowed only eight (8) days to be used for taking data; although the effort was made daily. Finally after 30% of the desired data had been collected, theft of equipment, lack of a stable platform in heavy seas, poor ocean visibility, and shortage of personnel dictated a change to

TABLE II

Control Sounds and Words

Phoneme	Word	Class of Phoneme
/AE/	AT	Vowel
/ε/	MET	Vowel
/P/	PAY	Voiceless Stop
/G/	GO	Voiced Stop
/ʃ/	SHOULD	Voiceless Fricative
/F/	FIRST	Voiceless Fricative
/TH/	THIN	Voiceless Fricative
/V/	VOTE	Voiced Fricative

TABLE III

Sample Campbell List (P-7)

DOLLS	AIM	COOK	GIVE
NUTS	CHAIR	AID	
BOOK	SMOOTH	AM	
JUMP	ARM	ACE	
ART	POOR	UP	
CARVE	HIGH	YES	
OR	SKIN	THREE	
ARE	LIE	HURT	

a quiet, fresh water environment having a stable platform. Two underwater communicators, the Hydro-Products 811 and the Bendix PQC-2 were used to transmit the data to an AN/WQC-1A communication set. Characteristics of this set appear in Table IV. This receiver fed a Precision Instruments DA 6200 tape recorder. The data was recorded at 37.5 ips and stored on magnetic tape. All recording used frequency modulation to allow for subsequent linear frequency translation by changing tape recorder speed.

A. ANECHOIC POOL

The anechoic pool located in Room 025 of Spanagel Hall, Naval Postgraduate School, was utilized to provide the noise free environment. The diver was dressed in normal diving equipment as shown in Fig. 5 and 6. Initially the diver spoke into free space as the control sounds, control words, and two Campbell word lists were recorded. Then an oral cavity was placed in position and the procedure repeated. A block diagram of the test setup is shown in Fig. 7.

B. OCEAN RANGE

The ocean range is located in Monterey Bay 1500 meters off Del Monte Beach. The location is shown in Fig. 8. The underwater portion of this range is still located as indicated. The range is located in 100 feet of water over a flat sandy bottom to minimize bottom reflection interference. Fig. 9 and 10 show the range layout. A diver platform is located at the main buoy to allow orientation of the diver toward the receiving transducer. This orienting device prevents signal loss due to attenuation by the lung cavities, wet suits, and other items of diver dress. Figure 11 shows a diver in the platform ready to

TABLE IV

Characteristics of AN/WQC-1A Communication Set

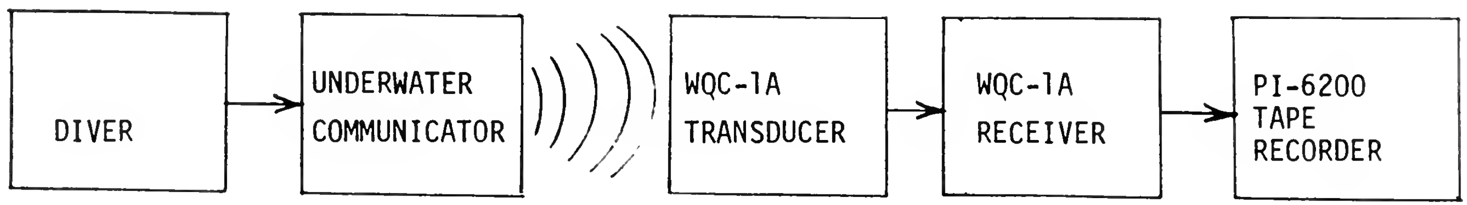
OPERATION	AM/SSB-SC
FREQUENCY	8.3 - 11 KHz
MODES	Receive, Transmit, Tone
TRANSDUCER	Omnidirectional w/o Baffle 15 db down 60°-330° w/ Baffle
BANDWIDTH	300-3000 Hz

Diver in Anechoic Pool
Without Oral Cavity

FIGURE 5

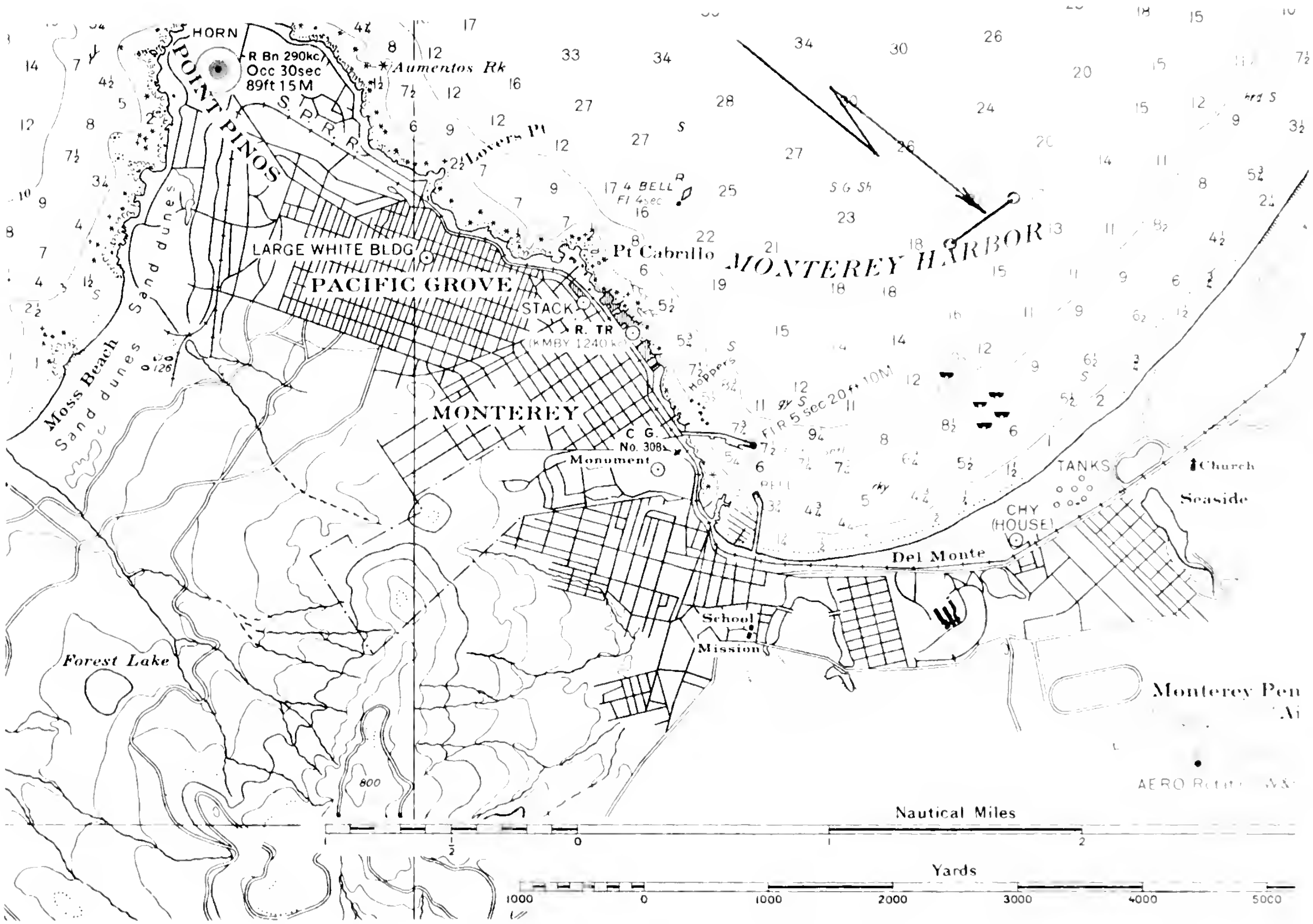
Diver in Anechoic Pool
With Oral Cavity

FIGURE 6



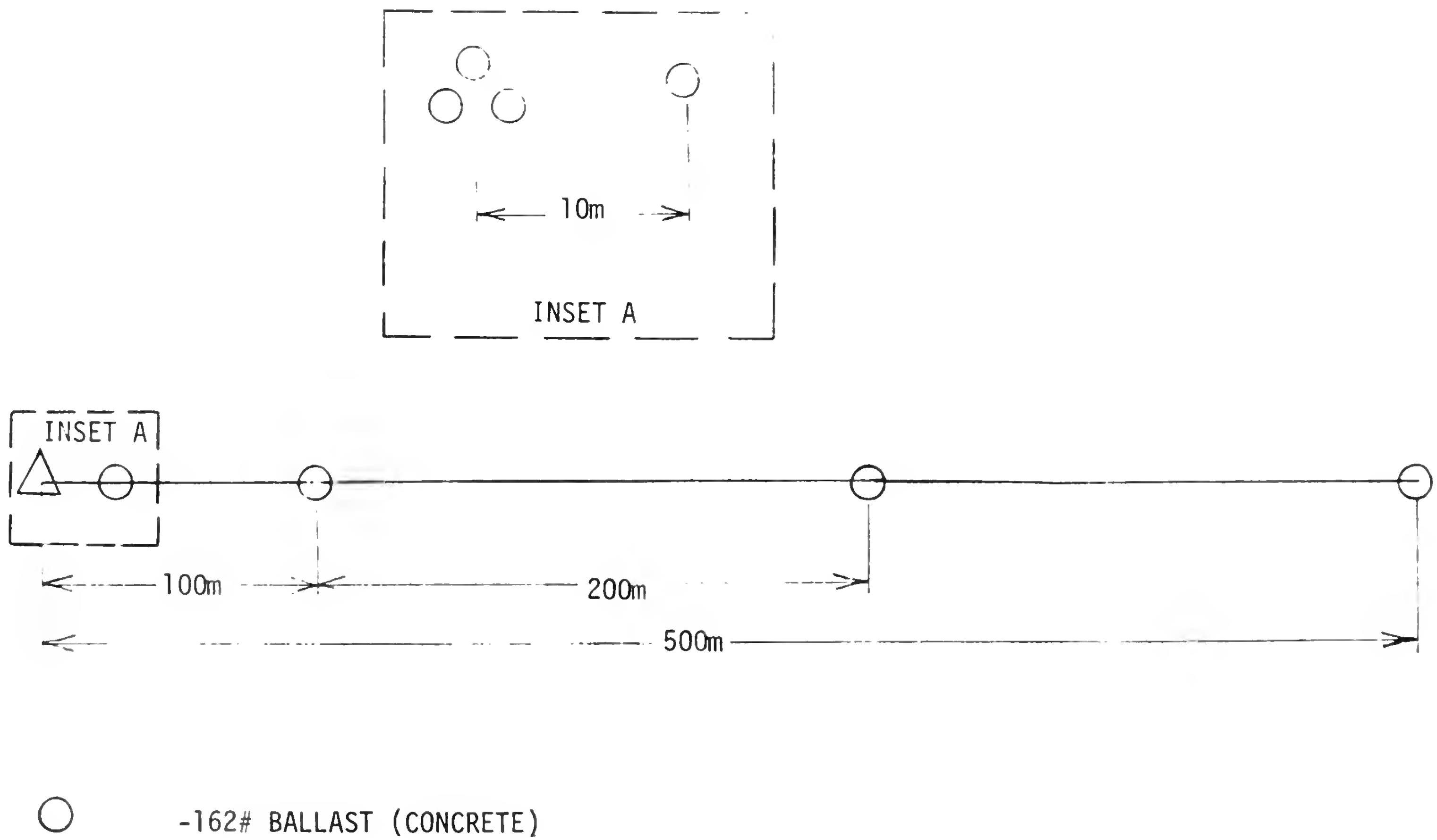
Block Diagram of Test Setup in
Anechoic Pool

FIGURE 7



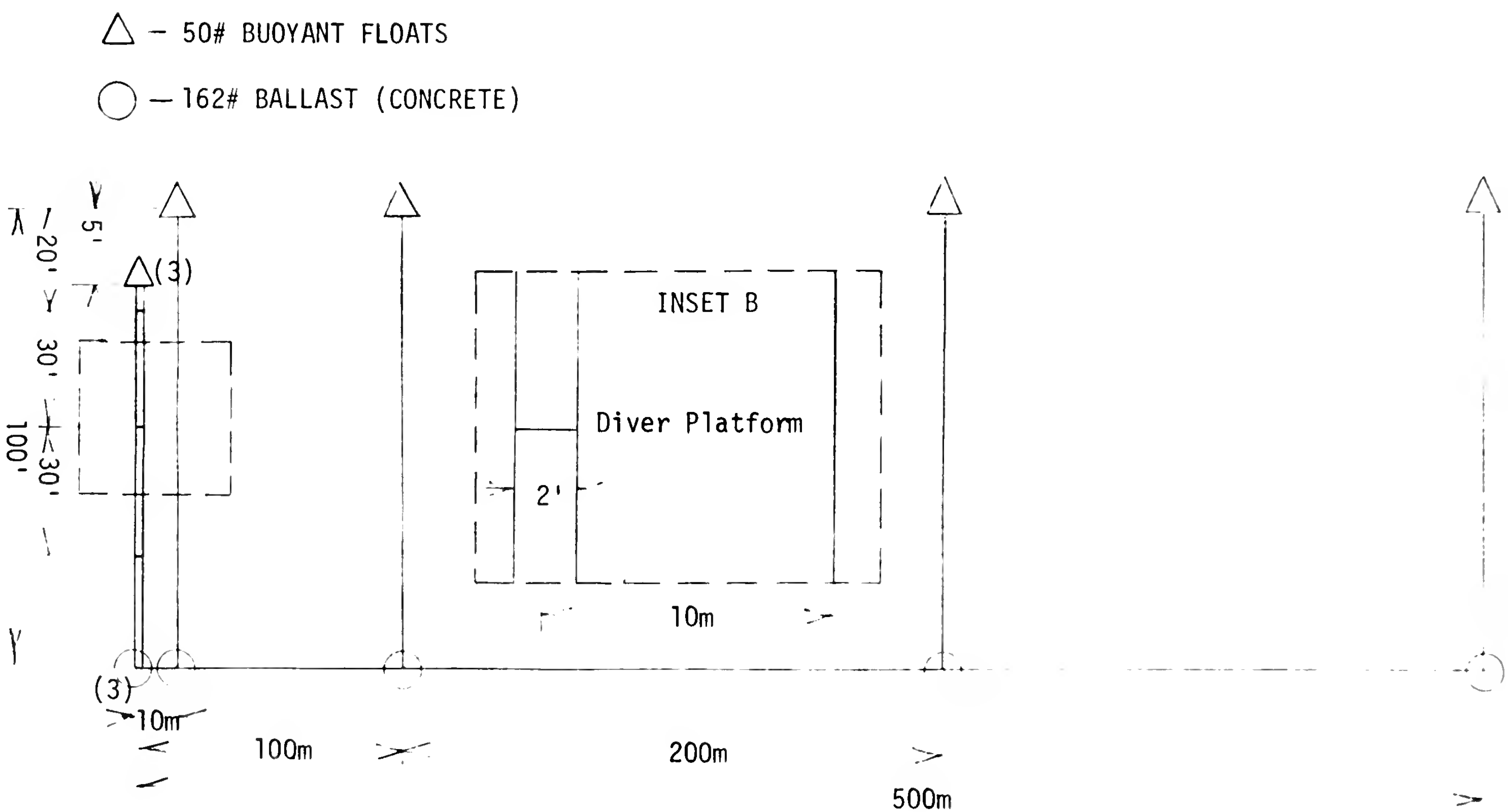
Map of Ocean Range Location

FIGURE 8



Ocean Range, Top View

FIGURE 9



Ocean Range, Side View

FIGURE 10

Diver in Orientation Platform

FIGURE 11

communicate. Figure 12 is a block diagram of the test set-up. Note that the tape recorder is mounted on an unstable platform (the boat) and must remain within short range (about 15 meters) of the desired buoy range marker. Table V indicates the desired conditions for recording data.

C. FRESH WATER RANGE

The fresh water range was set up in San Vincente Reservoir, San Diego County, California. Figure 13 shows the site. It differs from the previously described ocean range in four ways.

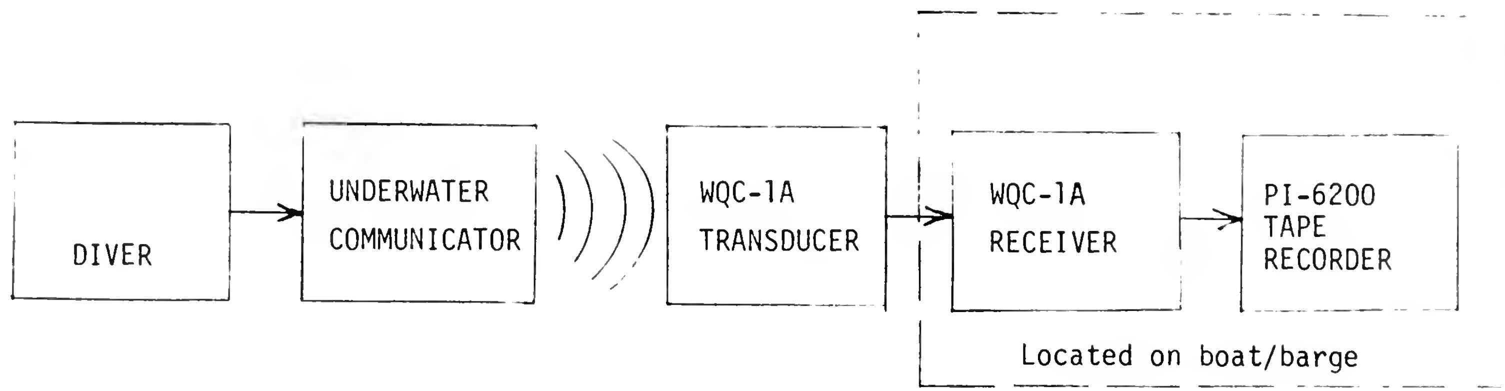
1. The granite walls and concrete dam face made receiving transducer orientation critical to prevent side reflections and minimize reverberation. This was accomplished by suspending the transducer on one line and using two guidelines to insure exact positioning.

2. The tape recorder was placed on a stable barge and the diver moved to different locations, thus reversing the roles in the ocean.

3. Quiet, fresh water replaced the noisy ocean environment.

4. Diver orientation was accomplished using a compass.

Table V shows the desired conditions for data collection. Eight days were required for movement to the reservoir, set up on the barge, collection of data and return to Monterey, California.



Block Diagram of Ocean Test Setup

FIGURE 12

San Vicente Reservoir Test Site

FIGURE 13

TABLE V

Recording Conditions for Ocean and Fresh Water Ranges

RANGE	DIVER SET DEPTH	BOND PQC-2	BOND HP 811	MYATT PQC-2	MYATT HP 811
10	20	X Y	Y	X Y	Y
10	50	X Y	Y	X Y	Y
10	80	X Y	Y	X Y	Y
100	20	X Y	Y	X Y	Y
100	50	X Y	Y	X Y	Y
100	80	X Y	Y	X Y	Y
300	50	Y	Y	Y	Y
500	50	Y	Y	Y	Y

X — DATA RECORDED IN OCEAN RANGE

Y — DATA RECORDED IN FRESH-WATER RANGE

III. DATA PROCESSING

The instruments available at the Naval Postgraduate School for the processing of the data include the IBM 360/67 computer, the SDS 9300 medium sized computer with a hybrid interface to a COMCOR 5000 analog computer.

A. ANALOG TO DIGITAL CONVERSION

The method of processing the data depended largely on the capabilities of the hybrid computer setup of the COMCOR 5000 and the SDS 9300. In order to digitize the data from the tapes, the sampling rate of the analog to digital (A-D) converter had to be at least twice as high as the largest frequency of interest (the Nyquist frequency). Unfortunately, the present write scheme of the SDS 9300 limits the sample rate to 2500 Hertz. This virtually eliminates the use of the A-D converter for the entire audioband. It was determined that the frequencies of interest for this experiment slightly exceed the passbands of the two communicators, i.e., 300-3000 hertz; a passband of 200-4000 hertz was used.

The sampling frequency dilemna was circumvented by utilizing the features of the FM tape recorder. The PI 6200 instrumentation tape recorder will record at three speeds - 37.5 ips, 3.75 ips, and .375 ips. The center frequency of the carrier is governed by the choice of the tape speed - 50,000 hertz, 5,000 hertz and 500 hertz, respectively. Suitable lowpass filters are switched into the circuitry for these speeds. At 37.5 ips the tape recorder input yields 0 db gain from dc to 10,000 hertz, with a signal to noise ratio of 42 db. By playing back at 3.75 ips, the discriminator frequency is divided by a factor ten, as are the signal frequencies. This linear frequency translation allows a sample rate of 800 hertz to cover the translated audioband.

The choice of parameters for sampling the data was hampered by the necessity for trade-off in frequency and time resolution, often referred to as the analogy to Heisenbergs "uncertainty principle" [13]. Ignoring statistical stability and those terms discussed in references on power spectra, the frequency resolution, Δf, is defined as:

$$\Delta f = f_s/N$$

where f_s = sampling frequency

N = number of samples

The dilemna is obvious. The larger the number of samples, the smaller the separation between spectral lines, hence, greater resolution. The larger N also extends the time duration the signal must be sampled, hence a loss of time resolution. To prevent aliasing, the sampling frequency must be at least twice the highest frequency of the signal. The larger sampling frequency reduces the frequency resolution.

The solution was to investigate the data for both time and frequency resolution. The time resolution was increased by using a lower number of samples for calculating an "evolutionary spectrum" and a larger number of samples for a spectrum of high resolution.

The sample rate was selected to be 2048 hertz. This frequency was selected because the FFT algorithm is most efficient when f_s is some number that is a power of two. It is more than five times the highest frequency of interest; i.e., 400 hertz (4,000 hertz translated by a factor of ten).

The number of samples taken must necessarily be a number that is a power of two for the FFT algorithm. Thus, the number of samples taken for each record was 8192. This allowed the signal to be sampled for a period of 0.4 seconds. The maximum resolution possible is 0.25 hertz, translated by the factor of ten to 2.5 hertz, real time.

To sample the data, the tape recorder output was filtered by a Krohn-hite adjustable bandpass filter, Model 310 A. The characteristics of this filter are shown in Table VI. The frequency limits were 20 and 400 hertz. The signal was then amplified by a factor of 30, and digitized with the clock rate controlled by the Wavetek 116 oscillator. The block diagram of Fig. 14 shows the sampling setup.

The signal was monitored by audio means and visually on an oscilloscope to determine the proper time to commence digitizing. Herein lies a fundamental weakness in this experiment; it will be discussed more fully in the analysis section. Briefly, no control was used to ensure that the sampling process started at the same instant the phoneme commenced.

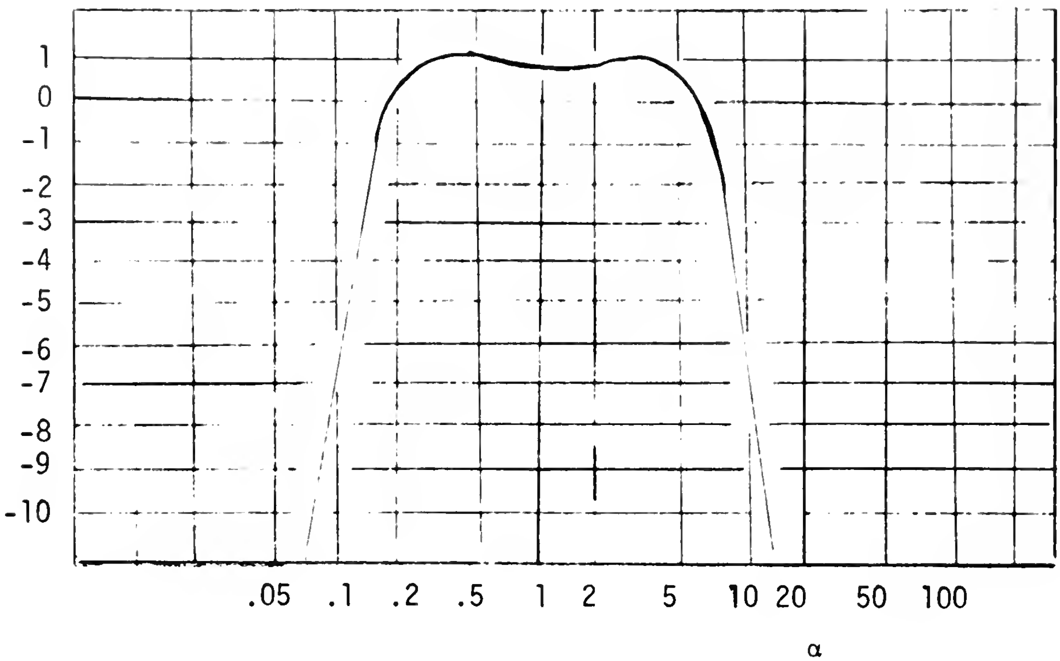
A vast amount of data was recorded for analysis. With time becoming a factor, it was necessary to limit the number of phonemes examined.

The phoneme selected was |AE|, the vowel sound in "at" - a front vowel. The same phoneme was sampled for the following conditions - each communicator at a range of ten meters and depth of 80 feet, and each communicator in the anechoic pool with and without oral cavity.

The sampled data from the SDS 9300 was placed on magnetic tape. The output was seven track in 24 bit octal format. The sample routine was an in-house routine developed by former NPGS students [14,15].

TABLE VI

Characteristics of Model 310A Krohn-hite Filter



$$\alpha = \frac{f}{90}$$

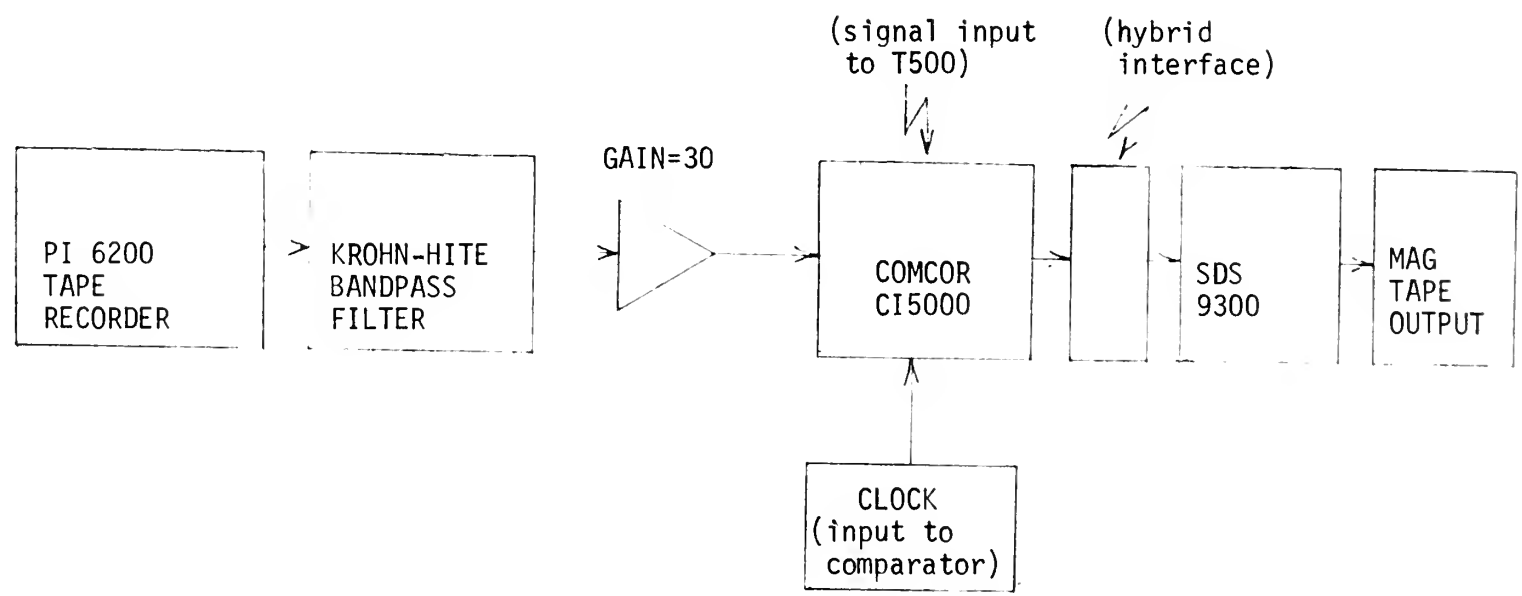
KROHN-HITE MODEL 310C

NORMALIZED FILTER RESPONSE

Attenuation: 24 db/octave below -12db point

Input Characteristics: 6 Megohms in parallel with 50 μμf
5 volts rms maximum

Output Characteristics: 200 ohms, 12 milliwatts



Sampling Block Diagram

FIGURE 14

B. INTERFACING

The output of the SDS 9300 is written in seven track. As is the case with all computers when trying to mate with the IBM 360, the interface is not as smoothly refined as might be expected. The seven track tape must be converted to nine tracks and the 24 bit, octal format converted to 32 bit, hexadecimal format. The program required is listed in Appendix B.

The tape drives of the IBM 360 required perseverance on the part of the operators. The tape from the SDS 9300 was accepted roughly one out of five tries.

No attempt was made to store this data on a direct access device. However,this seems desirable to conserve computer time.

C. DIGITAL PROCESSING

The data was stored on magnetic tape in 32 bit words. The subroutine Harm, part of the Standard Scientific Package for the IBM 360/67, was used to calculate the estimates for the Fourier coefficients of the signal. This routine is based on the Cooley-Tukey algorithm.

As previously mentioned, the entire 8192 data points could be used to compute a spectrum of high frequency resolution but with time uncertainties. Alternately, only a portion of the data points could be used with decreased frequency resolution but with increased certainty in the time domain.

The techniques complement each other. The technique chosen computed a) a power spectrum using the entire 8192 points, and, b) to assist in th analysis of this spectrum,computed a transient spectrum for the first

2048 points. The two spectra are titled "spectrum" and "transient spectrum." The spectrum has a resolution of 2.5 hertz. The transient spectrum has a frequency resolution of ten hertz.

The option exists whether or not a smoothing function is necessary. The need for a smoothing function might be described by mentally visualizing the side lobes of the sample function. These side lobes cause "leakage" from one sample to another. A smoothing function such as the now familiar "Hanning spectral window" reduces the side lobe leakage from $(w-w')^{-1}$ to $(w-w')^{-3}$ [16,17]. Whether or not the smoothing function is desirable depends on the type of signal. The literature says that noise-like signals with relatively flat spectrum are not suitable for smoothing functions, but when the signal is non-noise-like and the position of signal energy in the frequency domain is more important than the amount of energy, then a smoothing function is appropriate. The "Hanning window" was used. The estimates of the Fourier coefficients were operated on as follows (for N coefficients),

$$A(k) = -\tfrac{1}{4}\,a(k+1) - \tfrac{1}{4}\,a(k-1) + \tfrac{1}{2}\,a(k)$$

$$B(k) = -\tfrac{1}{4}\,b(k+1) - \tfrac{1}{4}\,b(k-1) + \tfrac{1}{2}\,b(k)$$

$$A(0) = \tfrac{1}{2}(a(0) - a(1))$$

$$B(0) = 0 \qquad k = 1,2,3,4$$

At the expense of a reduction in "statistical stability", this window was used; its time domain equivalent being,

$$X(j) = \tfrac{1}{2}\,(1-\cos(2\ j/n))\ x(j) \qquad j = 0,1,2,\ldots,n-1$$

A second smoothing operation was performed after the coefficients of the power spectrum were computed.

$$P(k)^2 = A(k)^2 + B(k)^2 \qquad k = 0,1,2,\ldots,N/2$$

$$P(k) = \tfrac{1}{4}(p(k+1) + p(k-1)) + \tfrac{1}{2}\,p(k) \quad k = 0,1,2,\ldots,N/2$$

This is again the familiar "Hanning window" [17]. Since the main concern is that the estimates are appropriate to the nominal frequencies, the linearly hanned spectrum suits its purpose.

The power spectra were plotted on a logarithmic scale, allowing a forced scale suitable for the entire dynamic range of data points.

IV. ANALYSIS

The phoneme selected for study was not an arbitrary choice. The noise-like sounds such as fricatives or stops do not lend themselves to Fourier analysis [18]. The vowel sounds, however, have a steady state quality, with slowly changing transient spectra. The front vowel, /AE/ as in "at", demonstrates a well defined spectrum with three important formant frequencies. The literature shows general agreement on the values of these formants:

$$F_1 = 660 \text{ hertz}$$

$$F_2 = 1720 \text{ hertz}$$

$$F_3 = 2410 \text{ hertz} \quad [19]$$

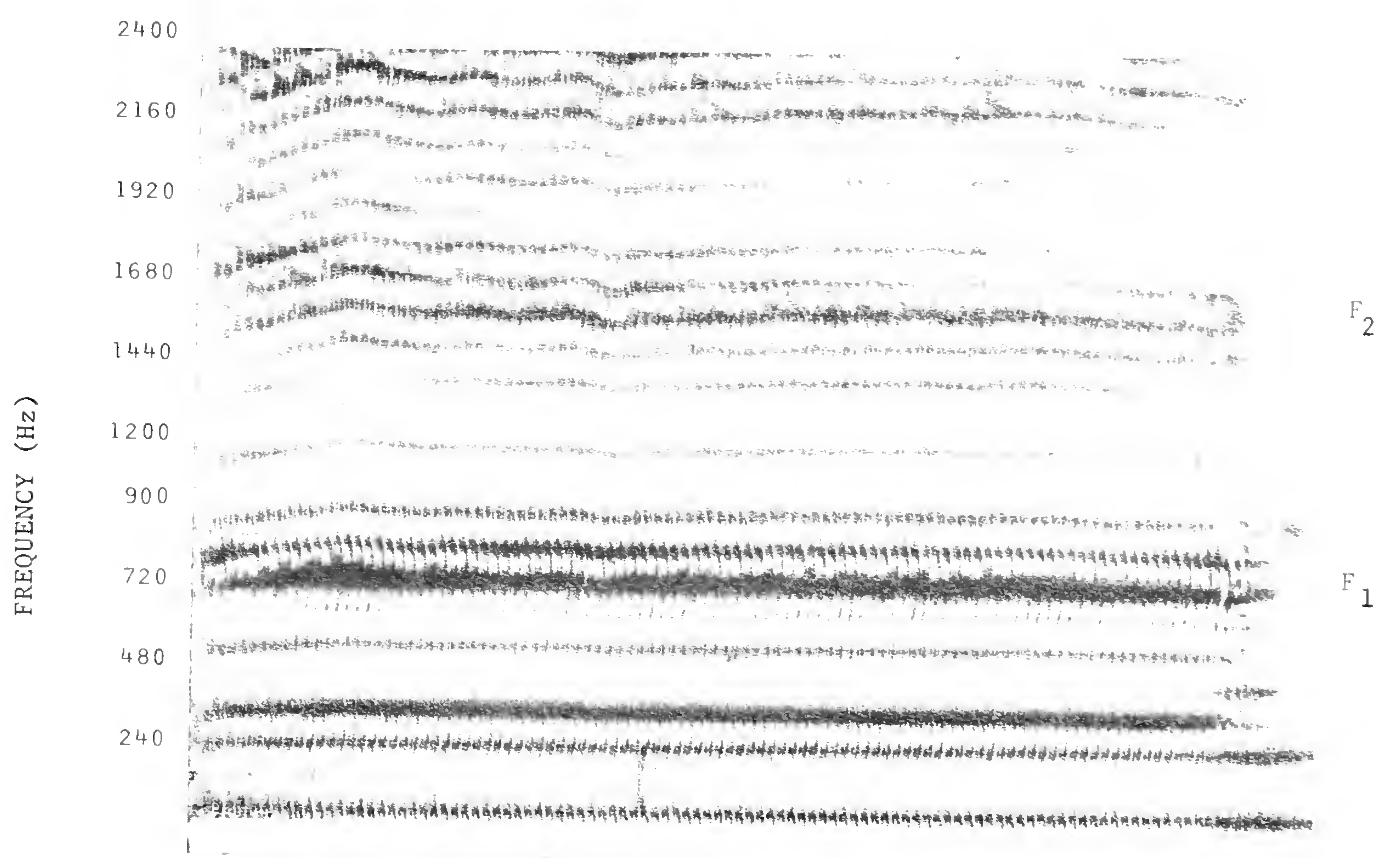
Associated with these formants are bandwidths of 60, 100, and 120 hertz, respectively. One reference shows the formants to be somewhat lower.

$$F_1 = 579 \text{ hertz}$$

$$F_2 = 1691 \text{ hertz} \quad [20]$$

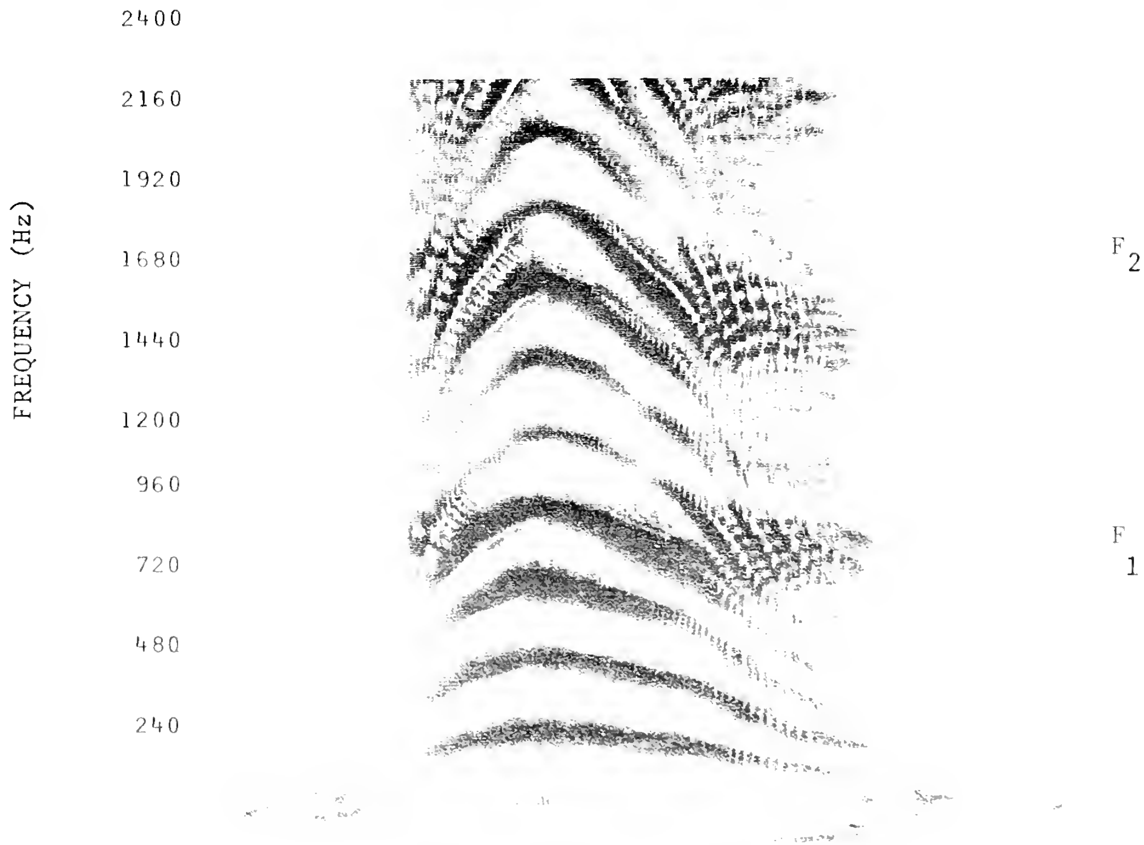
A. FORCED AND NORMAL SPEECH

In order to establish some sort of reference, a diver's voice was recorded on a Kay Missilyzer spectograph. Figure 15 shows the front vowel /AE/ recorded in an anechoic chamber. An Altec 21BR150-2 microphone was used and the diver spoke in a normal voice. Figure 16 is the same sound spoken in a loud, forceful manner, just as if the diver were using an underwater communicator. The requirement for the forced speech to increase intelligibility has been documented [1].



Spectograph of Front Vowel /AE/ (Normal Speech)

FIGURE 15



Spectograph of Front Vowel /AE/ (Forced Speech)

FIGURE 16

The formant frequencies for the normal speech are shown with arrows in Figure 15. They generally agree with the values cited previously. Figure 16 shows the formants greatly increased. The effort to increase the level of sound has caused the diver to lift the pitch of his voice in an unnatural manner. The spectrum is not constant but rises rapidly to a maximum value for each formant and then slowly decays. This change in the spectrum is of major importance. Further verification of this difference in the forced and normal speech power spectral estimates is borne out by Figures 17 (a-c) and 18 (a-c). The normal speech power spectral output closely approximates a line spectra where shifts can be easily measured.

B. FREQUENCY TRANSLATION

To check the frequency translation of the PI-6200 tape recorder, a 384 hertz tone was recorded at 37.5 ips and sampled at the rate of 2048 hertz. The power spectrum is shown in Figure 19. Figure 20 shows the power spectrum when the tape recorder speed is slowed to 3.75 ips. The expected change in frequency by a factor of ten occurred.

C. POWER SPECTRAL ESTIMATES

1. Without Oral Cavity

A sample spectrum for the recording made using the communicators without oral cavity is shown in Figure 21(a-c). The spectrum is cluttered with power at frequencies in the neighborhood of the formants. This clutter is sufficient to prevent accurate measurement of the formants; but should be anticipated in light of the observations made concerning the rapid changes in the spectrum in Figure 16. This

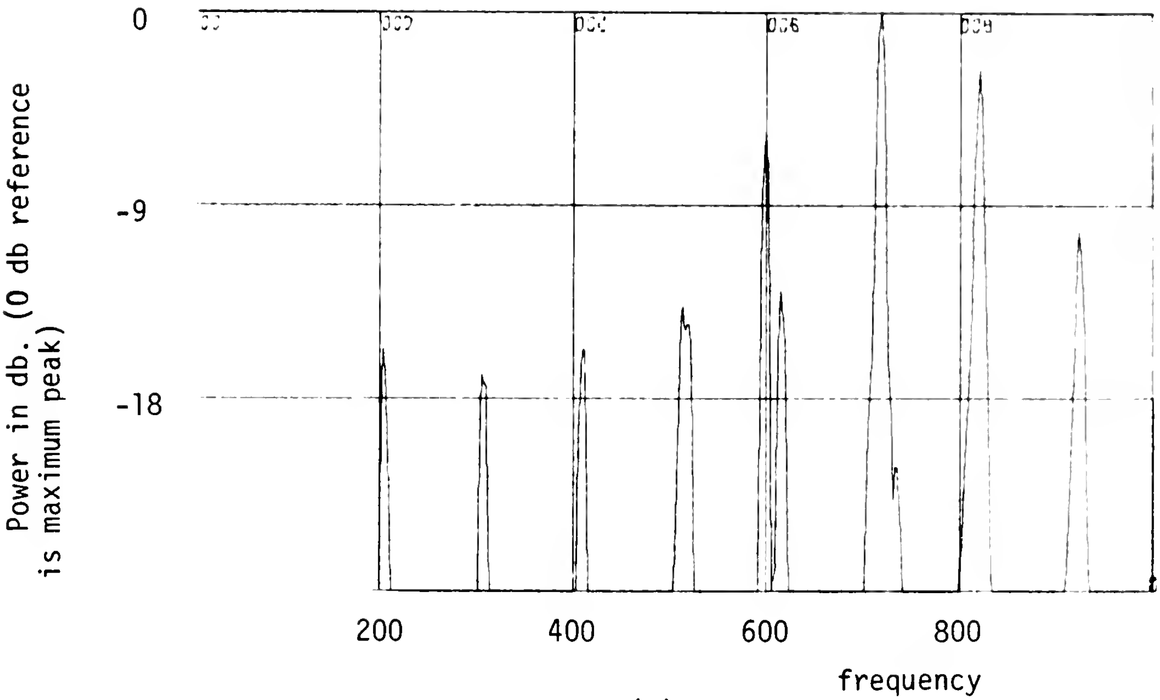


FIGURE 17(a)

Normal Speech Spectrum of /AE/, 0-1000 Hz.

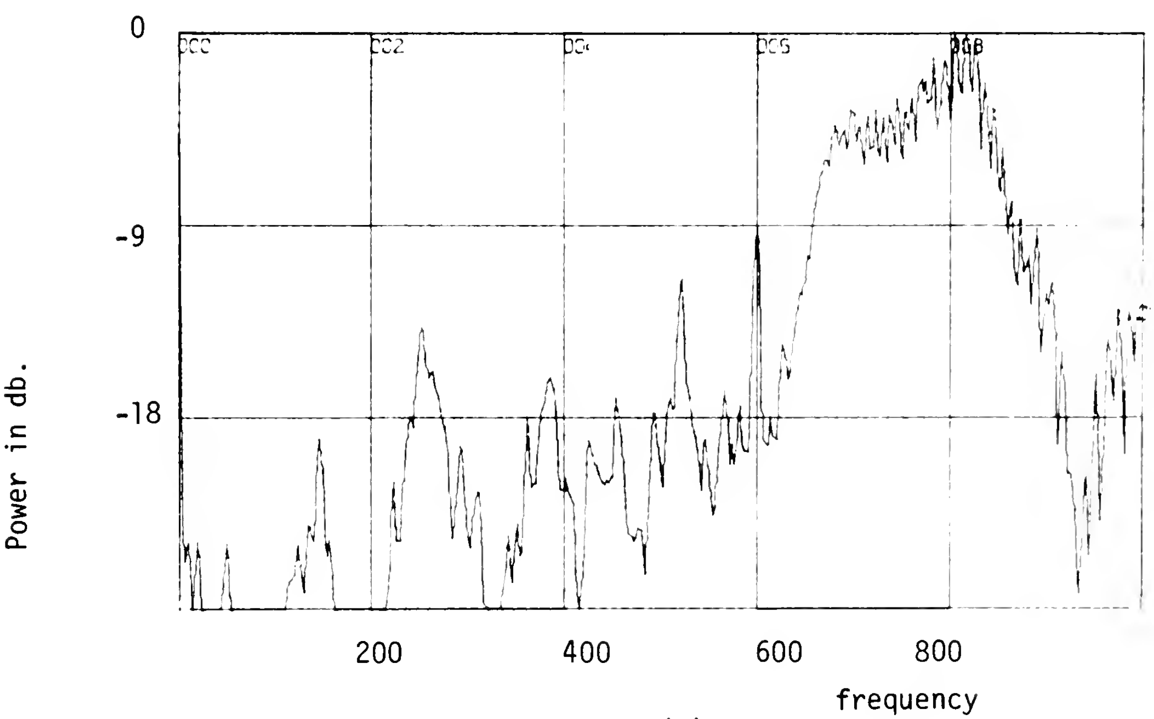


FIGURE 18(a)

Forced Speech Spectrum of /AE/, 0-1000 Hz.

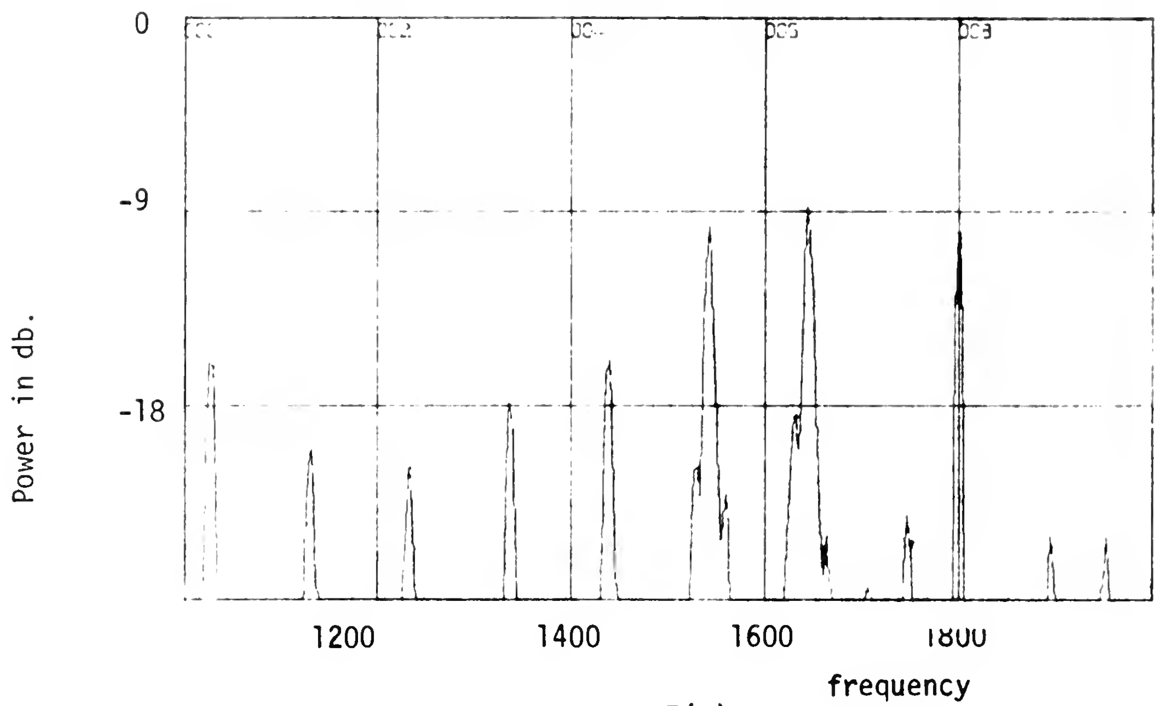


FIGURE 17(b)

Normal Speech Spectrum of /AE/, 1-2 KHz.

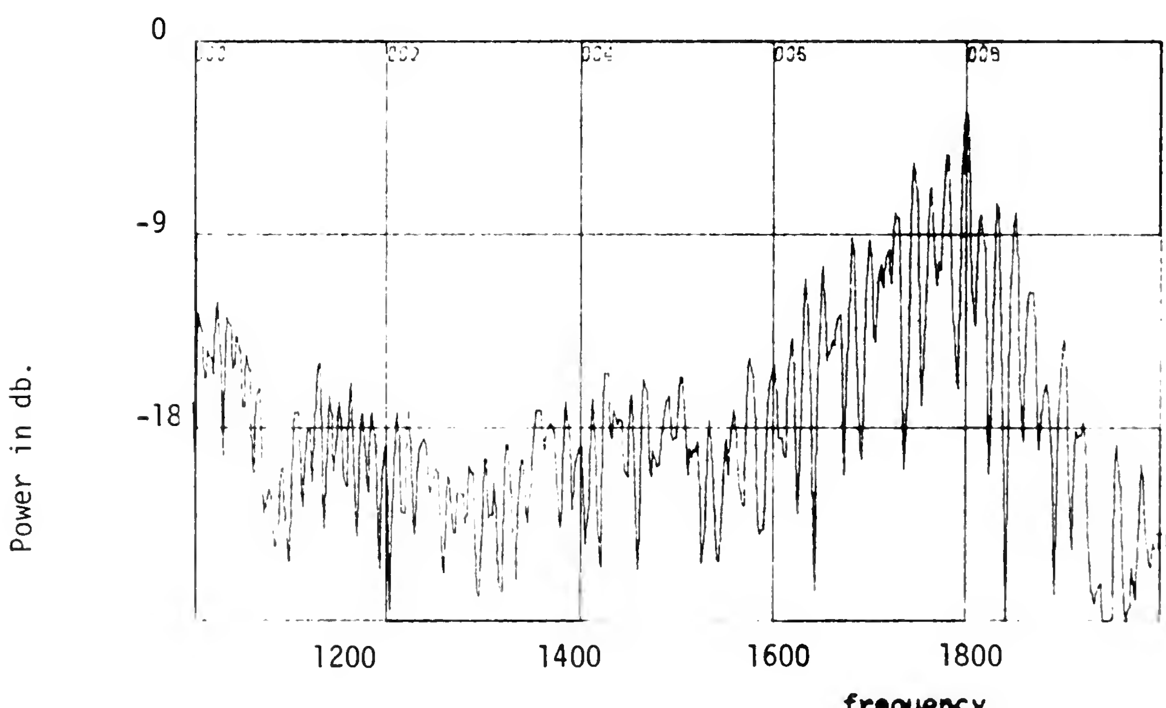


FIGURE 18(b)

Forced Speech Spectrum of /AE/, 1-2 KHz.

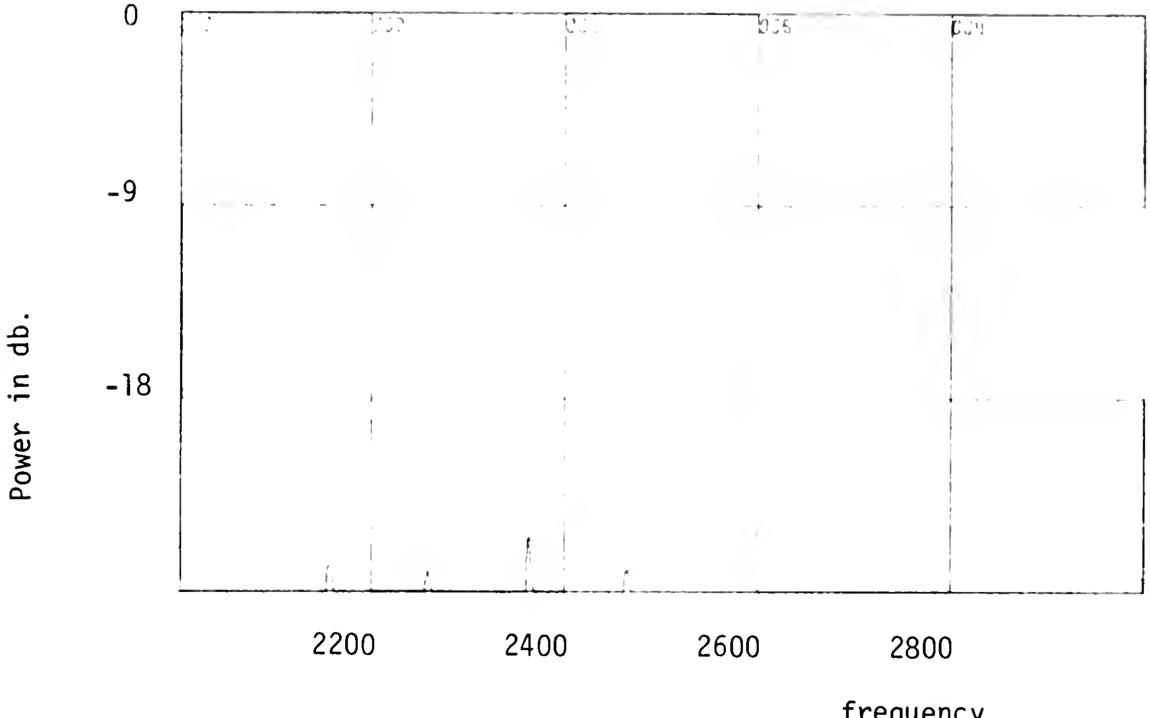


FIGURE 17(c)

Normal Speech Spectrum of /AE/, 2-3 KHz.

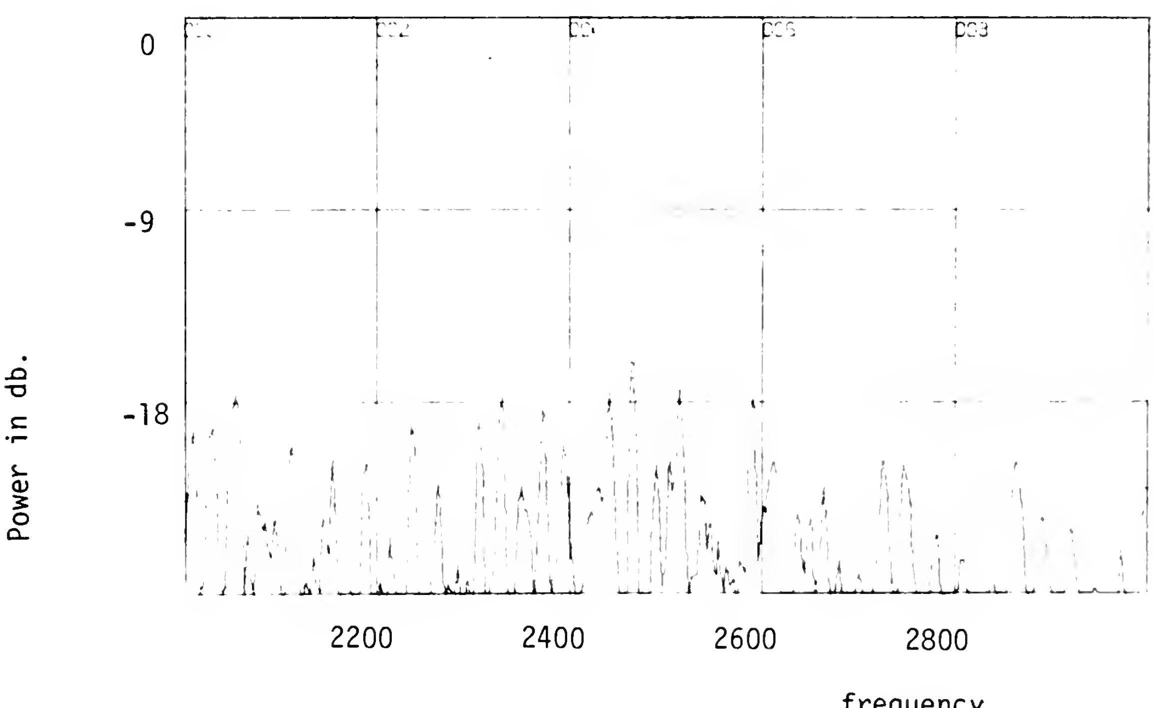


FIGURE 18(c)

Forced Speech Spectrum of /AE/, 2-3 KHz.

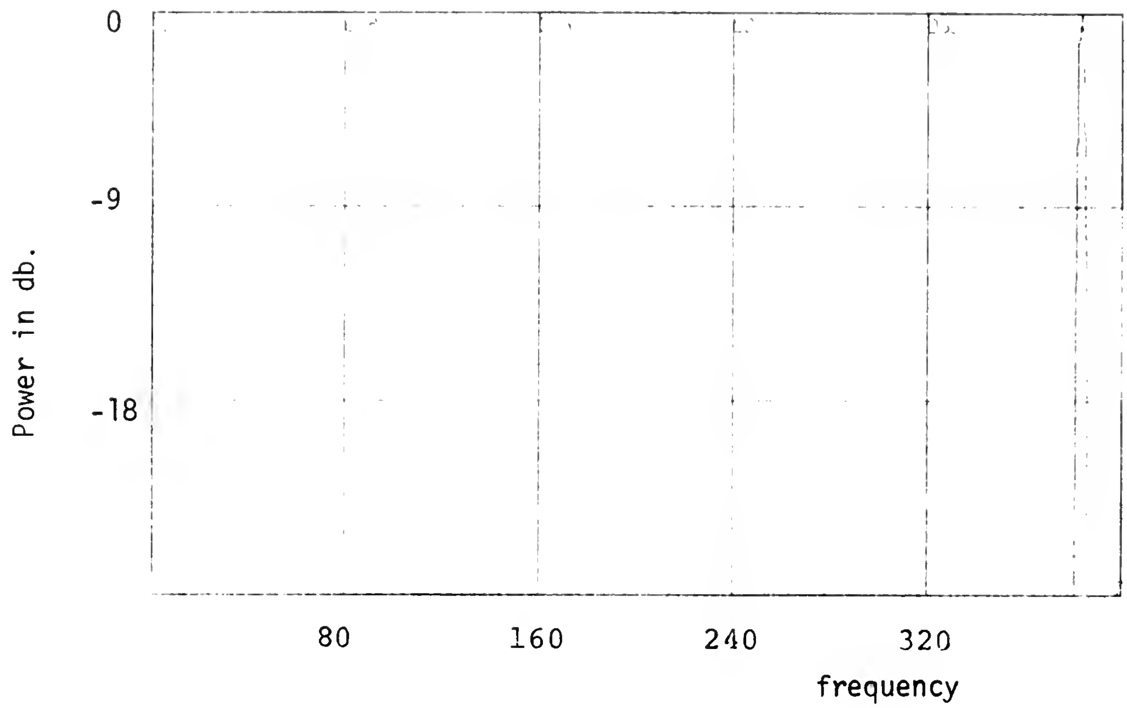


FIGURE 19

384 Hz. Tone at 37.5 ips

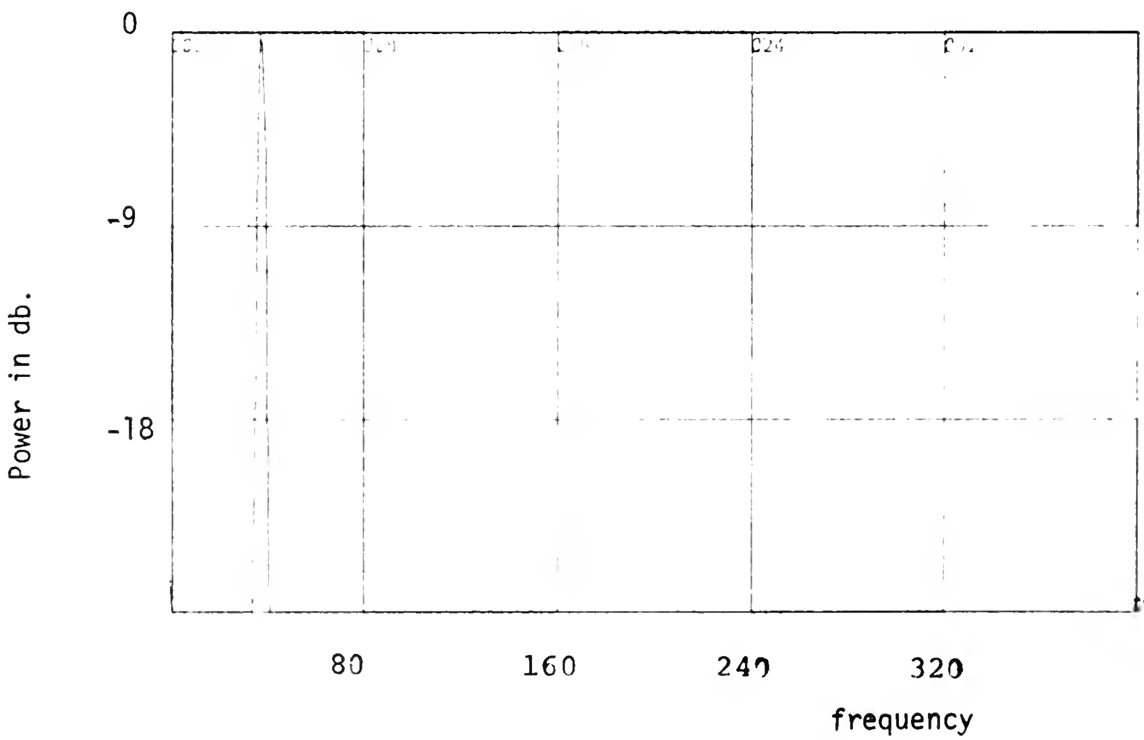


FIGURE 20

384 Hz. Tone at 3.75 ips

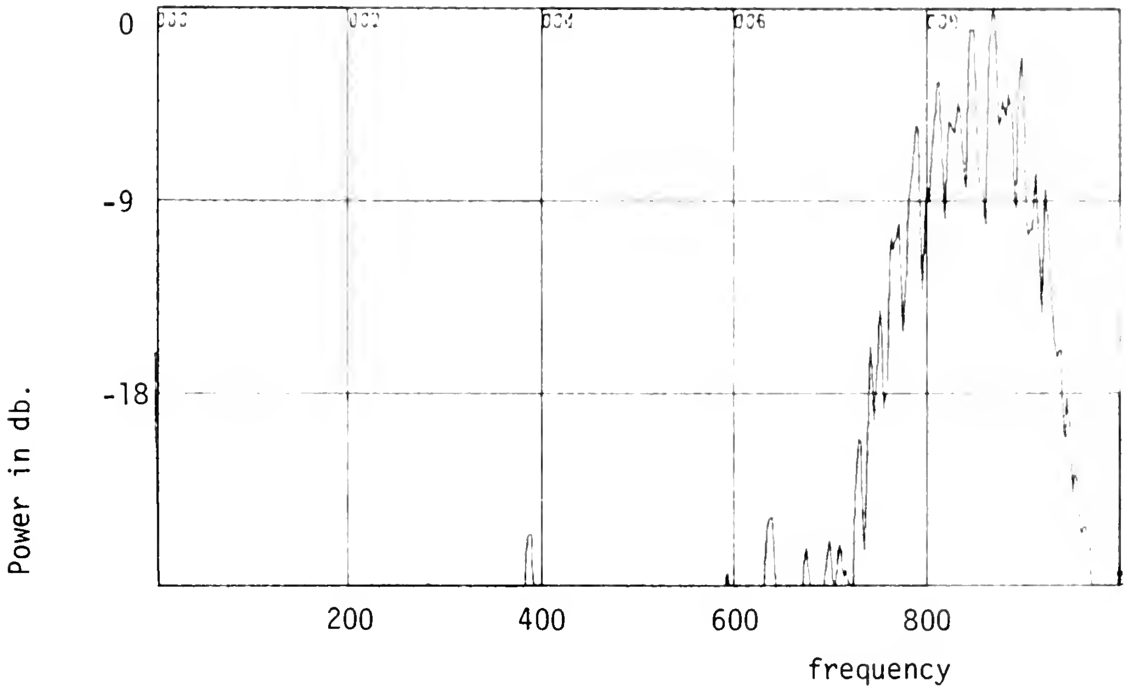


FIGURE 21(a)

Phoneme /AE/, Without Oral Cavity, 0-1 KHz.

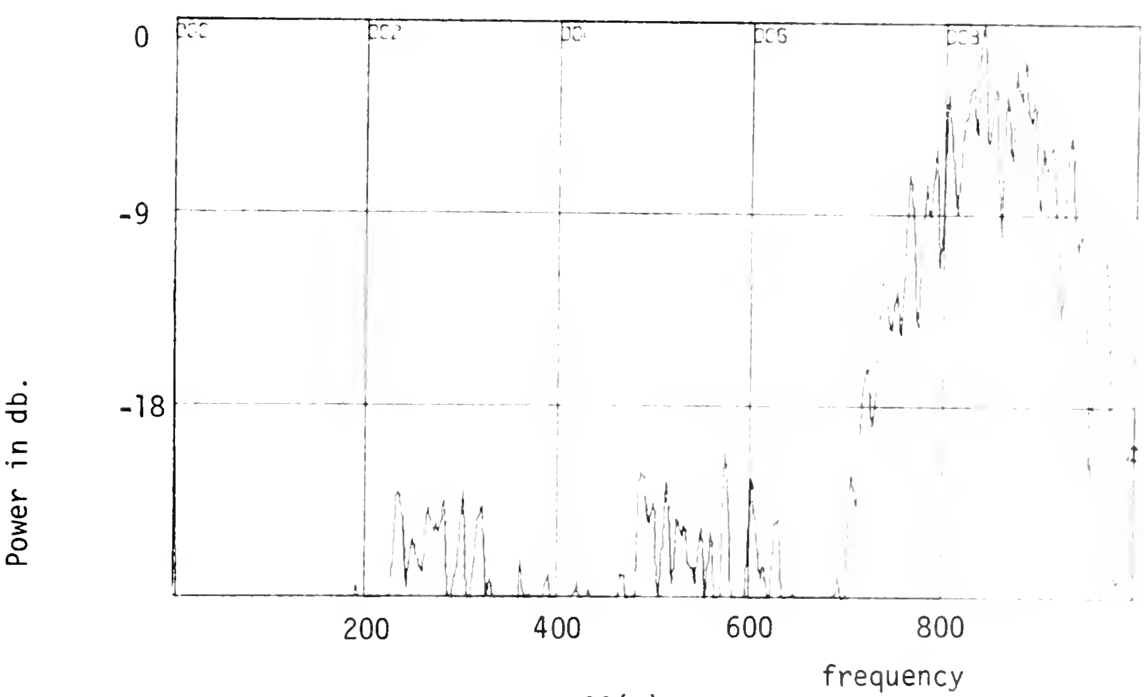


FIGURE 22(a)

Phoneme/AE/, With Oral Cavity, 0-1 KHz.

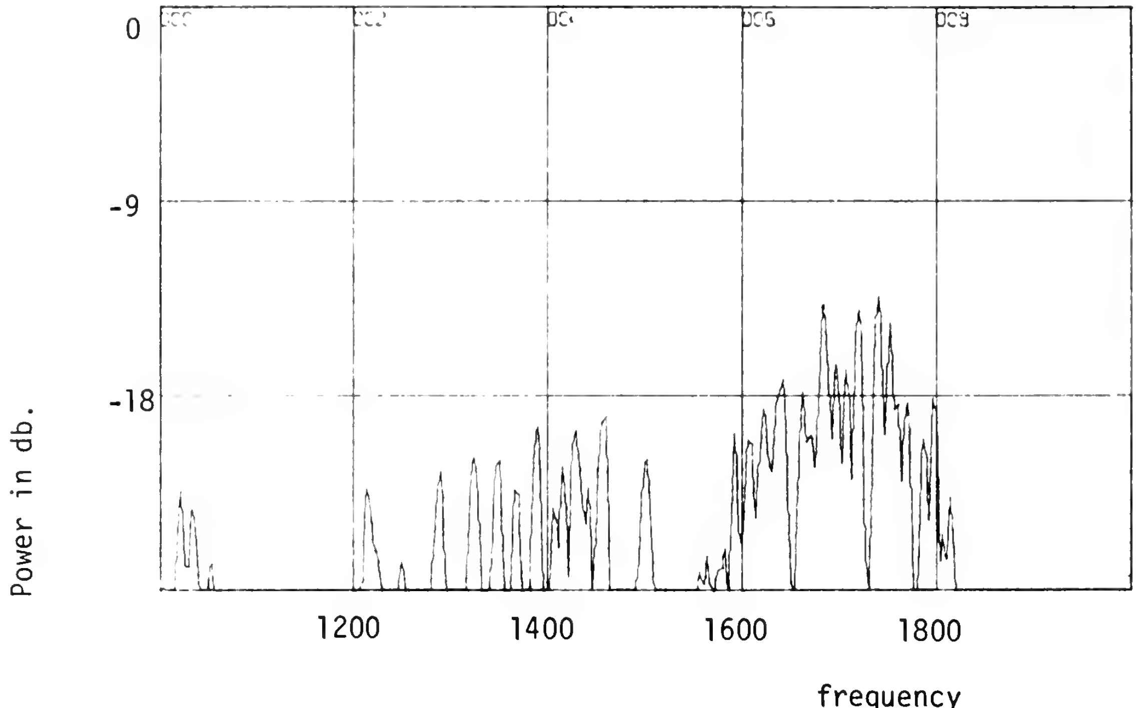


FIG. 21(b)

Phoneme /AE/, Without Oral Cavity, 1-2 KHz.

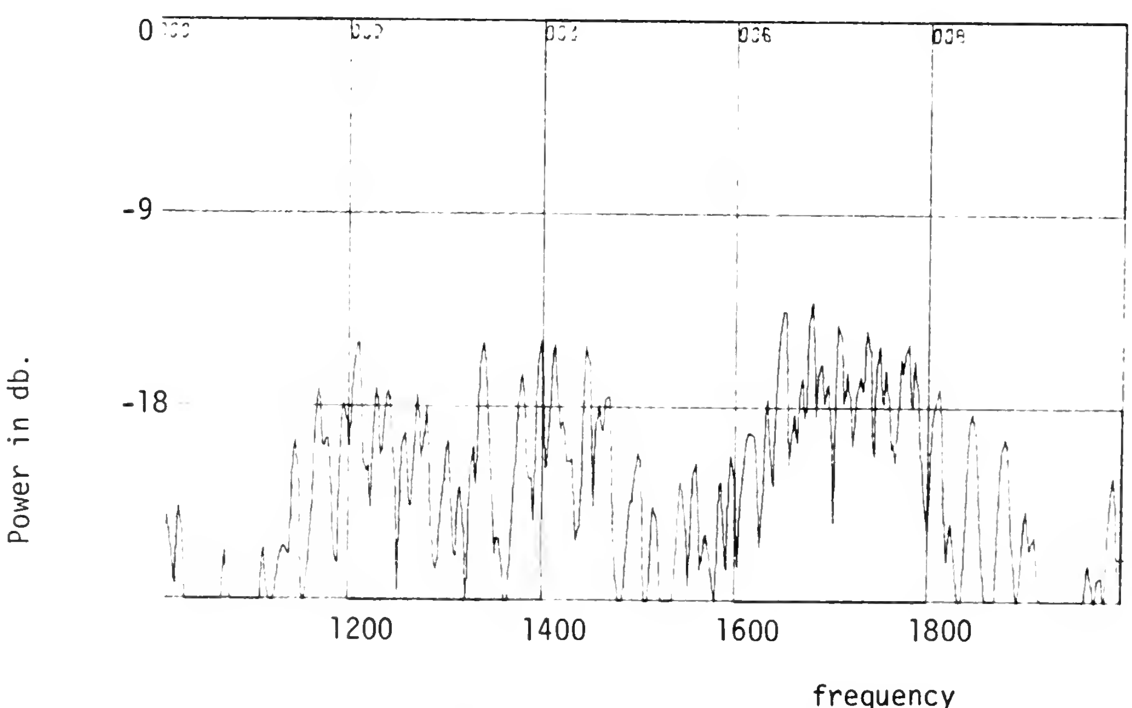


FIG. 22(b)

Phoneme /AE/, With Oral Cavity, 1-2 KHz.

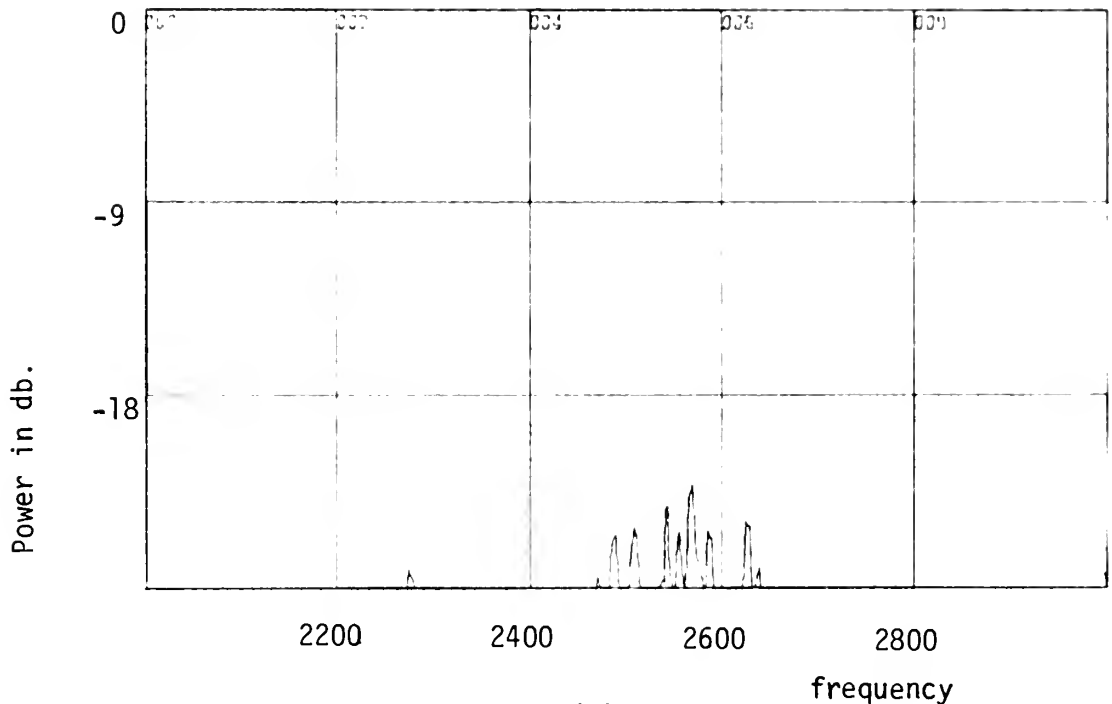


FIG. 21(c)

Phoneme Without Oral Cavity, 2-3 KHz

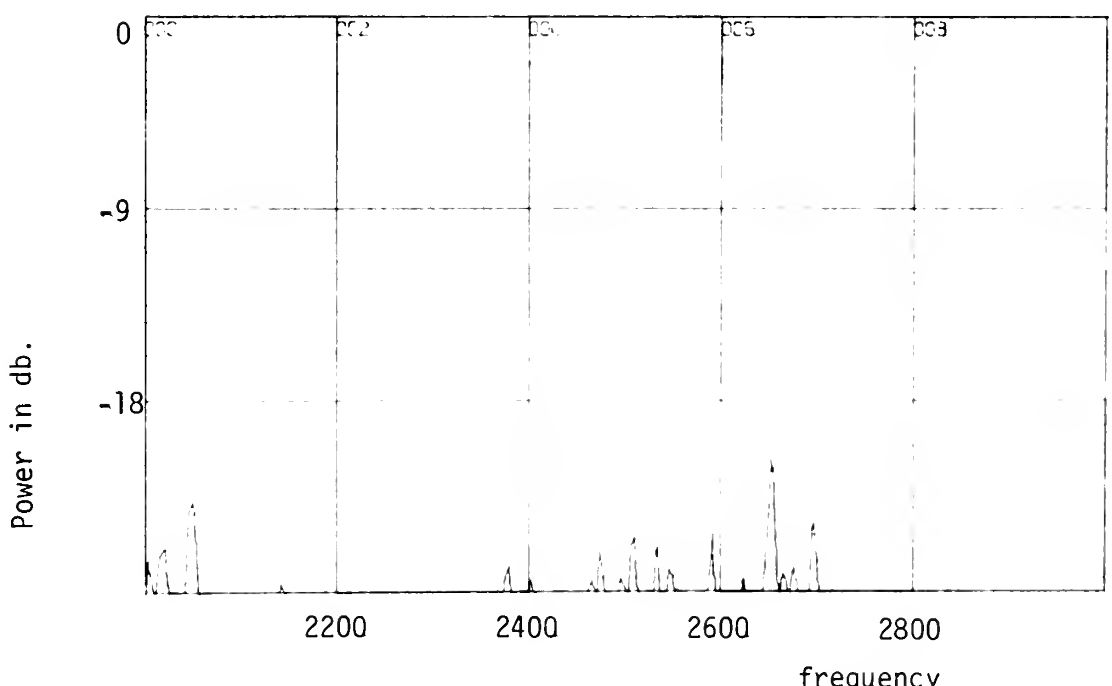


FIG. 22(c)

Phoneme /AE/, With Oral Cavity, 2-3 KHz.

spectrum was estimated using 8192 sample points. Theoretically, the frequency resolution is 2.5 hertz. In reality, only a rough estimate can be made to determine the formants within 50 hertz. Blaming the poor results on the manner in which the diver forced the sound, the next step was to compute an 'evolutionary spectrum', (transient spectrum). Using the first 2048 sample points, the theoretical resolution is ten hertz. The transient spectrum for the first 2048 points is shown in Figure 23a-c. This covers a signal for 0.1 second. The obvious blunder is that no technique was used to ensure that the sampling process began precisely at the beginning of the sound. This is critical in the transient spectrum, in that an additional uncertainty has been introduced. That is; does the sampling start before the spectrum reaches the highest values in frequency? If the signal had been the phoneme recorded in the normal voice, this would not have been critical since the spectrum is flat for a reasonable period of time.

Since the data was already recorded, it was not possible to record on an adjacent channel, any form of signal to coincide with the start of the sound so that the sample process could have been automatically triggered. The alternative was to simultaneously monitor the signal by audio and visual means and manually trigger the sampling process. Obviously this technique has flaws and the exact point of entry into the data stream is uncertain.

Even with the ambiguities, the fact that the formants are shifted up in frequency due to the forced sound level is apparent. The first

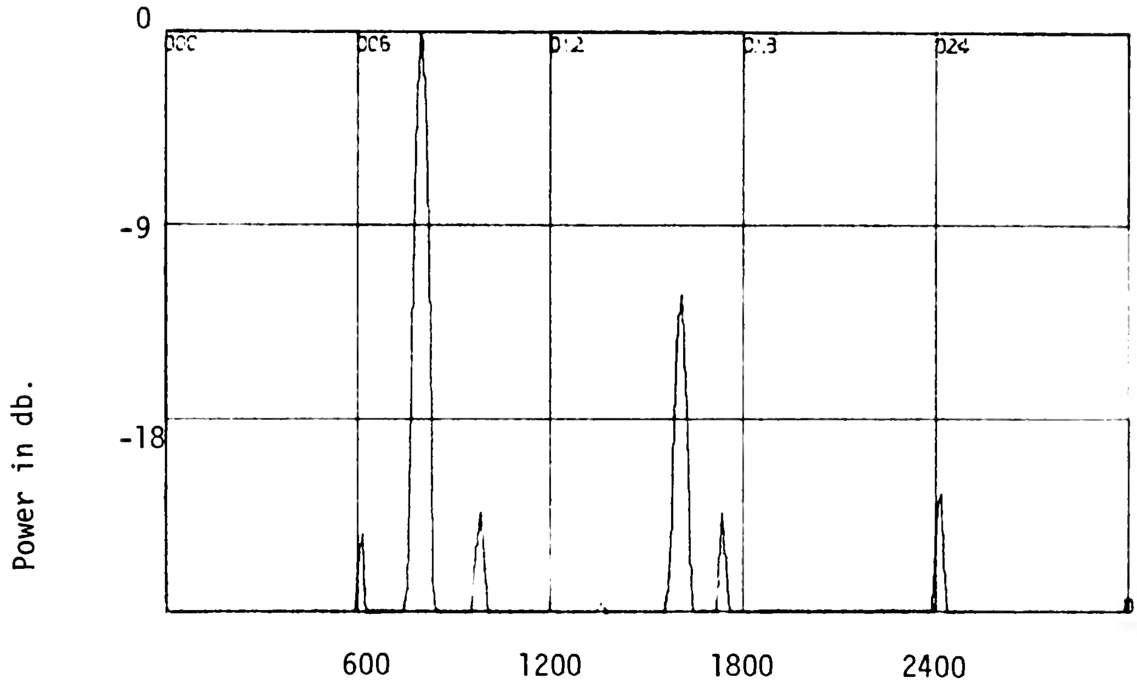
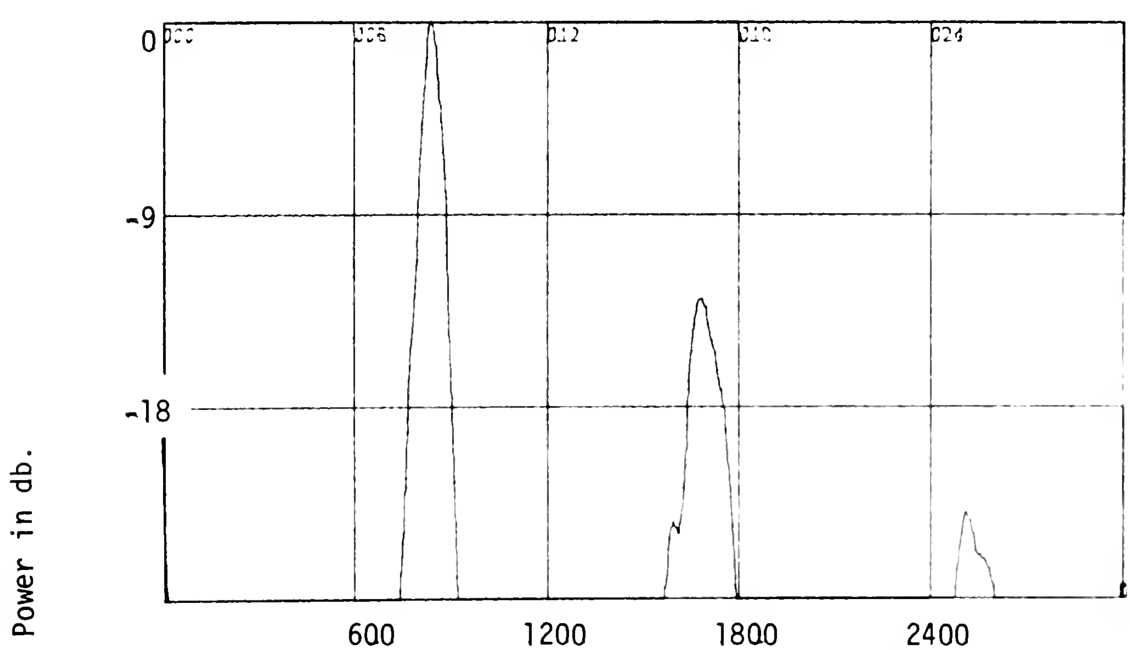


FIG. 23 frequency

Transient Spectrum, Without Muzzle, 0-3 KHz



frequency

FIG, 24

Transient Spectrum, With Muzzle, 0-3 KHz.

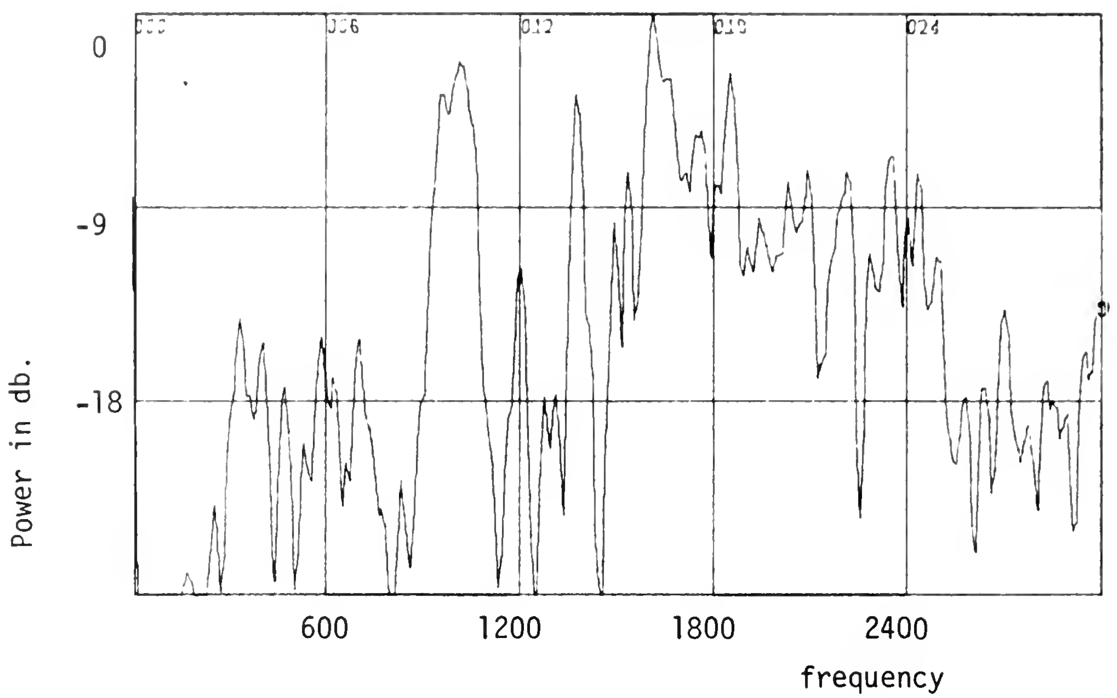


FIG. 25

Transient Spectrum, 80 feet, 0-3 KHz.

formant can be estimated at 840 hertz, in agreement with Figure 15. The second formant is centered at 1720 hertz and the third at 2550 hertz. Bandwidth is extremely hard to estimate.

2. With Oral Cavity

A sample spectrum for the recordings made using the communicators with the oral cavity is shown in Figure 20 (a-c). The literature concerned with small, non-radiating enclosures, such as the oral cavity used, states that in theory the effect of the muzzle is to raise the formants in frequency; however in reality, the distortion is much more complicated. "... The formants may become indistinct rather than shift in any direction..." [21]. These observations are partially borne out in the spectra referenced above. F_2 does appear to be indistinct as does F_3. F_1 contains the major concentration of power, the center being shifted down instead of up in frequency. The inability to massage the data to make concrete observations is regretable.

3. Increased Ambient Pressure With Oral Cavity

A sample spectrum for the recordings made at 80 feet is shown in Figure 26 (a-c). The effect of increased pressure is well documented [1,5,6]. The formants should shift upward in frequency in a non-linear fashion. "...the first formant is affected more than the second formant. It is also observed that the intensity level of voiced sounds increased with increasing ambient pressure but that typical noise sounds, such as fricatives, and the burst part of unvoiced stops display a drop in intensity relative to voiced sounds ..." [6]. This referenced source cites a theoretical expression for computing the upward shift in frequency of the first formant.

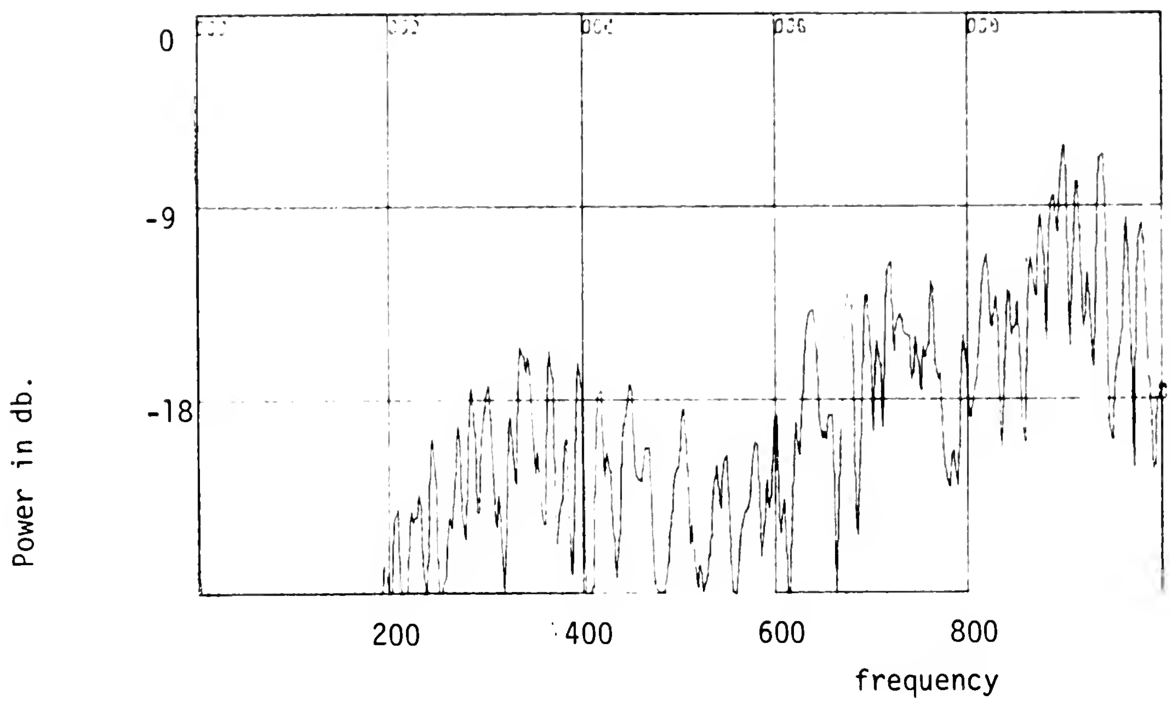


FIG. 26(a)

Phoneme /AE/, 80 feet, 0-1 KHz.

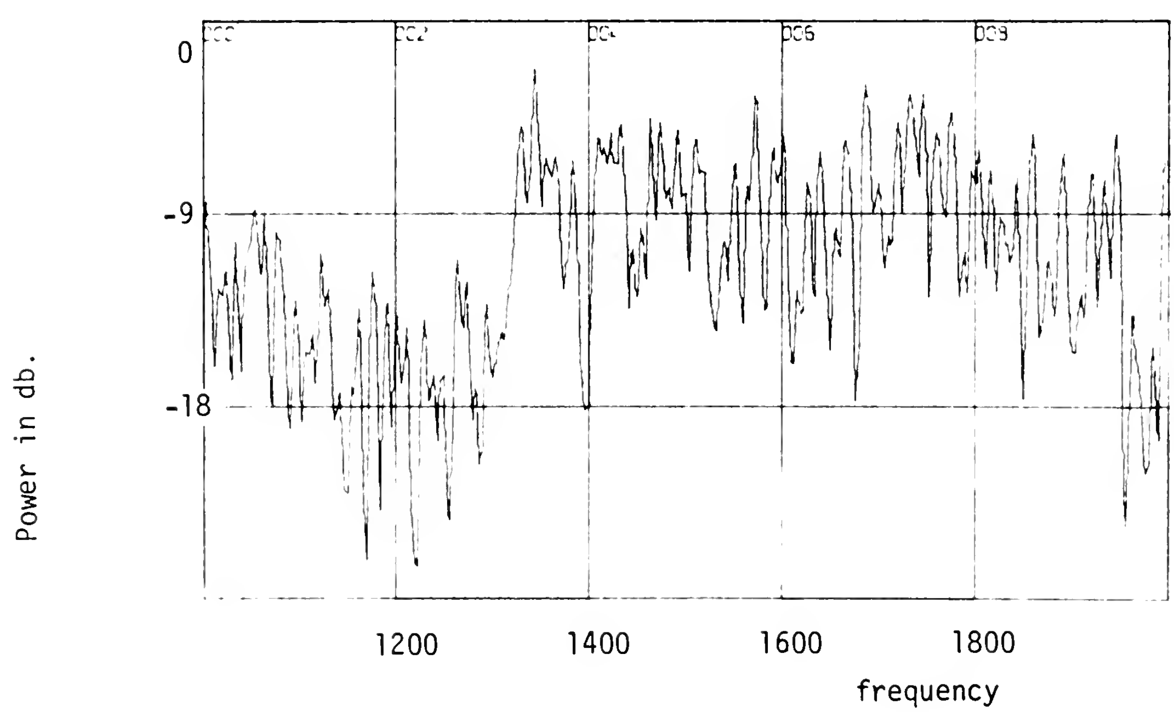


FIG. 26(b)

Phoneme /AE/, 80 feet, 1-2 KHz.

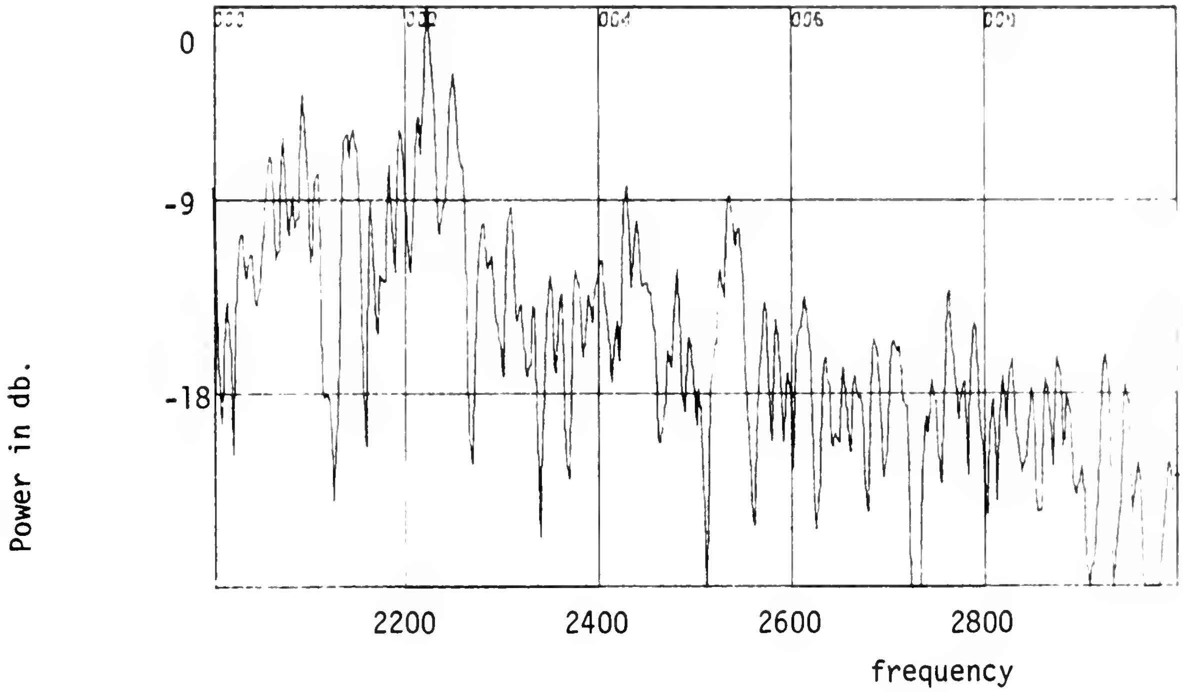


FIG. 26(c)

Phoneme /AE/, 80 feet, 2-3 KHz.

$$F_{1p}^{\;2} = F_{11}^{\;2} + F_{w1}^{\;2}\,(p-1)$$

where F_{1p} is the frequency of the first formant at p atmospheres pressure, F_{11} is the first formant at one atmosphere, and F_{w1} is the resonance frequency of the closed vocal tract at one atmosphere. Applying this expression to the situation under discussion, namely a depth of 80 feet, the theoretical first formant shift can be calculated. Assuming a resonant frequency of 170 hertz, and using the estimate of 840 hertz for F_{11}, with p = 3.42 atmospheres, F_{1p} is computed to be 878 hertz, a shift of 38 hertz. Examining the spectrum for 80 feet in Figure 26, the first formant can be roughly estimated to be 910 hertz, significantly larger than expected. The source of this shift is uncertain; it could be related to the increased effort of the diver, as well as to the increase in pressure. F_2 and F_3 are indistinct, the power distributed in many frequencies throughout the band.

V. CONCLUSIONS AND RECOMMENDATIONS

The original objective of isolating the distortion of the underwater communicators was not accomplished. Ths problem narrowed to developing a tool to examine the speech power spectrum in detail with high resolution. It is felt for non-noise like phonemes that the FFT is the tool available to the analyst to perform this high resolution analysis. Some of the final massaged output reinforces this belief. Line spectra can be the final result of speech wave forms if a number of the troubles encountered are corrected in subsequent research. To assist in future work a summary of the troubles encountered and recommendations for correction are submitted.

A. DATA COLLECTION

1. Equipment

The proper equipment must be available to support the research in all phases of data collection and remain available for recall during the processing phase. This affects a number of topics, range location, etc., but is best illustrated by the necessity for long term availability of the underwater communicators. In this experiment, the communicators were on short term loan and a number of experimental techniques were hastily contrived to support the data collection.

The proper diving equipment must be made available to support the collection effort. Diver comfort in cold water requires the addition of heated suits to prevent changes in speech due to loss of heat. The orientation problem, discussed later, calls for a helmet or hood type lantern. The type developed for Sea Lab III by Ben Saltzer of Naval Undersea Warfare Center, Pasadena, California is recommended. Short term loan of personal equipment and use of outdated equipment complicates the overall effort and reduces efficiency.

2. Tape Recorder

The tape recorder deserves special mention; the choice is critical. The PI 6200 was chosen because it could be adopted for dc power and provided the necessary linear frequency translation due to the slow sample rate. The PI 6200 tape recorder is delicate and required special handling, but is suitable for use at a fixed recording site.

3. Range Location

The range location is critical. The attempt to set up a range in Monterey Bay during the worst weather in years cost many man hours. The range was completed and used; however theft by local fishermen negated the effort. If Monterey Bay is to be used in the future (it should not!) the summer months provide the best ocean conditions, and notices should be put out to all fishermen asking them to cooperate in avoiding destruction of the range. The Monterey Harbormasters Office will assist in this respect. Time and effort can be saved if a better location is used, specifically, Stage I at NSDRL, Panama City, Florida.

4. Diver and Receiving Transducer Orientation

Diver orientation must be closely controlled as must the orientation of the receiving transducer. The lung cavities and wet suit cause the transmission of the communicators to be highly directional. If the diver is not oriented properly, the signal strength will be reduced; similiarly the WQC-1A transducer is directional, so must be oriented to provide maximum signal reception. In Monterey Bay the divers orientation was fixed by a platform allowing him to sit down. No attempt was made to orient the transducer in the ocean range. At San Vicente Reservoir the granite walls and concrete dam face required that the receiving transducer be oriented to preclude reflections. The diver was oriented using a wrist compass to maintain the proper line to the receiving transducer. The diver held a flashlight in one hand to assist in reading a word list in total blackness, while he clung to a safety line to control his depth and location. It is no wonder the result was fading due to diver drift from the proper diver/transducer line. A setup of the DICORS type would materially assist in solving the orientation problem [5].

5. Safety

Additional divers, preferably with a technical background, are needed for safety and convenience. Throughout this research, the divers violated normal safety standards by diving alone. This left one man at the surface to control and monitor the myriad of controls for recording. The ability of the topside man to monitor the data collection with confidence was hopeless.

6. Spurious Noises

The attenuation of noises associated with tape recorder heads, interference of 60 cycle nature, surface and bottom reflections must be constantly monitored. Data should be recorded only if optimum conditions exist.

7. Forced Speech

The data collection must support the data processing. The very fact that two independent projects were undertaken reduced the effectiveness of either project. The method of collecting data for intelligibility tests is vastly different than that to support the speech processing — the subject of this paper. Intelligibility tests are optimized when the diver raises the level of his voice and makes a determined effort to speak clearly. However, as mentioned repeatably, the task of isolating the problems of distortion by use of the power spectrum requires the diver to speak in the same manner throughout. The line spectrum desired is most nearly approached when the voice is articulated in a normal manner. This implies that instead of forcing the voice on the surface to agree with the technique used for clarity at depth, the voice must be used in a normal manner at each stage of the experiment, a reverse of what was done.

B. DATA PROCESSING

The processing technique must be thoroughly understood and be used successfully prior to the collection of the data. Lack of a hybrid interface and limited availability of the underwater communicators precluded these conditions being met.

1. Sample Rate

The hybrid computer facility should modify the write scheme to ensure that the A-D conversion rate is high enough to accommodate speech signals. The desired rate is at least ten kilohertz. The analog computer, the COMCOR 5000, is presently rated at 25 kilohertz; the limiting item is the digital computer, the SDS 9300. The limitation of 2.5 kilohertz forced the conversion of the data at a playback speed of 3.75 ips, one-tenth the record speed. In theory, no error is introduced but it is desirable to process at the record speed.

2. Filtering

Filtering of the signal prior to A-D conversion should be variable and of high quality. The Krohn-hite filter is suitable for such work and in many ways is superior to a fixed Bandpass filter using active devices. A voltage controlled bandpass active filter would also be suitable.

3. Interfacing

Converting the seven track, 24 bit tape to a nine track, 32 bit tape was a particular stumbling block, requiring long hours to overcome. Appendix B is included to assist in overcoming this problem in the future.

4. Output Format

The output from the IBM 360 can be displayed in a variety of ways; however, the CalComp Offline plotter is most suitable for analysis. A more efficient draw routine would make processing easier and save compute time and core memory.

APPENDIX A GLOSSARY

PHONEME— one of a finite (40) number of sounds that characterize speech. Six major subgroups (classes) are shown in Table VII.

FRICATIVE— a noise-like sound, primarily produced by blowing air over the teeth, as in "s", "sh", and "f."

STOP— a sound produced by holding the breath and exploding it in a burst of noise, as in "p", "t", and "k."

VOWEL— a steady state sound as in all common vowels.

GLIDE— a sound formed by continually moving the articulators (mouth, tongue, etc.), as in "l", "r", and "w."

NASAL— "m" and "n"-like sounds.

DIPTHONG— two or more vowels sounded together

PLOSIVE— a stop

FORMANT— the resonant frequencies of the vocal tract.

AFFRICATIVE— a phoneme beginning with a stop and ending with a fricative, as "ch" in the word "church."

PITCH— a qualitative dimension of hearing related to the highness or lowness of a tone and the sound intensity (with increased intensity, a relative change in pitch can occur with a change in frequency.)

TABLE I [21]

STOPS

/p/	pat	/b/	be
/t/	to	/d/	day
/k/	key	/g/	go

FRICATIVES

/h/	house	/v/	vest
/f/	fee	/ð/	then
/θ/	thin	/z/	zoo
/s/	see	/ʒ/	azure
/ʃ/	she		

NASALS

/m/	me
/n/	need
/ŋ/	song

GLIDES

/w/	we
/j/	you
/r/	read
/l/	let

VOWELS (including dipthongs)

/i/	eve	/AE/	at	/ɔi/	boy	/u/	foot
/I/	it	/a/	ask	/ɒ/	not	/u/	boot
/e/	hate	/ʌ/	up	/ɔ/	all	/ai/	I
/ɛ/	met	/ei/	say	/ə/	obey	/au/	out
		/ɑ/	father			/ou/	go

APPENDIX B

The purpose of the appendix is to give a brief but complete description of the mechanics involved in digitizing data and interfacing from the SDS9300 to the IBM 360/67.

A. ANALOG TO DIGITAL CONVERSION

The analog signal is the input to the Comcor CI 5000. For all signals, a bandpass or lowpass filter network is necessary to prevent aliasing, a major problem. The upper cutoff frequency should be the highest frequency of interest. The filtered signal is plugged into an amplifier input on the CI 5000 and amplified by the highest possible factor within the ± 100 volt limits of the operational amplifiers. This is required to yield the highest possible signal to noise ratio. The output of the amplifier is the input to Trunkline 500(T500) on the analog patchboard. This links the analog to the digital computer.

The sampling frequency can be selected up to and including 2500 hertz for the present digital magnetic tape write scheme. If the analog clock pulse is used, a selection of 10, 100, or 1000 hertz is available on the logic patchboard and should be connected into the Master Clock Input (MC IN). The output of the Master Clock (MC) is plugged into Trunkline 210 (T210) on the logic patchboard.

If a variable sampling frequency is desirable or if a rate different than what is available on the logic board is required, an oscillator can be plugged into a comparitor input, (say C03) on the analog patchboard. The output of the comparitor on the logic patchboard is plugged into Master Clock In (MC IN). The output of the Master Clock (MC) is connected to Trunkline 210, (T210).

The sampling program for the SDS 9300 is titled SAMPL, and was developed by Lt. Lynn Dorrian for the SDS 930 [14]. It was modified for the SDS 9300 by Lt. Jerry Post and Mr. R. Limes [15]. The program will sample one analog signal. The data is stored in buffers and written on magnetic tape from these buffers. It is designed to fill one buffer while the other buffer is dumping onto the magnetic tape. Since the write scheme for the magnetic tape is inherently slow, the sampling rate is presently limited to a maximum of 2500 Hz. If only two buffers of data is desirable, this rate can be slightly exceeded.

The procedure is to call the Subroutine SAMPLE(NREC,NSAMP,IBUF0, IBUF1,NTAPE,ITOG) where:

NREC = number of records/run.
NSAMP = number of samples/record.
IBUF0 = first data buffer.
IBUF1 = second data buffer
NTAPE = magnetic tape unit number
ITOG = 0 if IBUF0 is currently being filled.
= 1 if IBUF1 is currently being filled.

After the program has been accepted by the SDS 9300 and the parameters initialized, the digital computer waits for the analog computer to go to "compute." The compute switch can be manually triggered or triggered by a control signal.

Once the A/D routine commences, it continues until either the analog goes out of "compute" or the number of records of data has been taken. Insure after the proper number of records have been taken that the keyboard is put back into reset, if triggering manually. Control returns to the digital program. The digital program waits for instructions. If Sense Switch 6 is lighted, the program returns to initialization and waits for the analog to again go to compute to digitize a new run of data. If Sense Switch 5 is lighted, the magnetic tape is rewound to load point. If Sense Switch 4 is lighted, the End of File condition is entered on the tape. Sense Switch 3 will cause the line printer to print the last 2 records of data in a print plot routine.

Enclosure (1) is a diagram of the CI 5000 analog and logic patch-board configuration for the A/D conversion. Enclosure (2) is the program listing for the Fortran IV routine utilized.

Below are listed a number of hints to save effort and time.

1) Use a magnetic tape provided by the IBM 360/67 computer facility as it has a permanently marked load point with the correct spacing. The tape unit number is 1 and should be set on "automatic."

2) Ensure that digital program uses fresh cards, particularly the RTM BOOT card; a feed check on the card reader is often sensed as a program error.

3) Follow basic 9300 instructions. If the control console fails to start program feed, turn off the system and start over. This "immediate action battle drill" could save many hours.

4) Use multichannel oscilloscope on the analog console to monitor the pre-filtered signal, the post-filtered signal and the amplified signal. This will allow a visual check to ensure proper analog action.

5) Monitor hybrid interface console during digitizing to ensure that the system is working.

6) Make at least the first two and last two runs from a known analog signal (known frequency and amplitude). This is necessary to ensure the conversion to the IBM 360 is complete and correct. An additional suggestion is to insert a run of known data periodically throughout the data of interest to also allow a check of the conversion.

7) Read NPGS Computer Facility Technical Note 0211-64 prior to attempting A/D conversion.

B. CONVERSION TO THE IBM 360/67

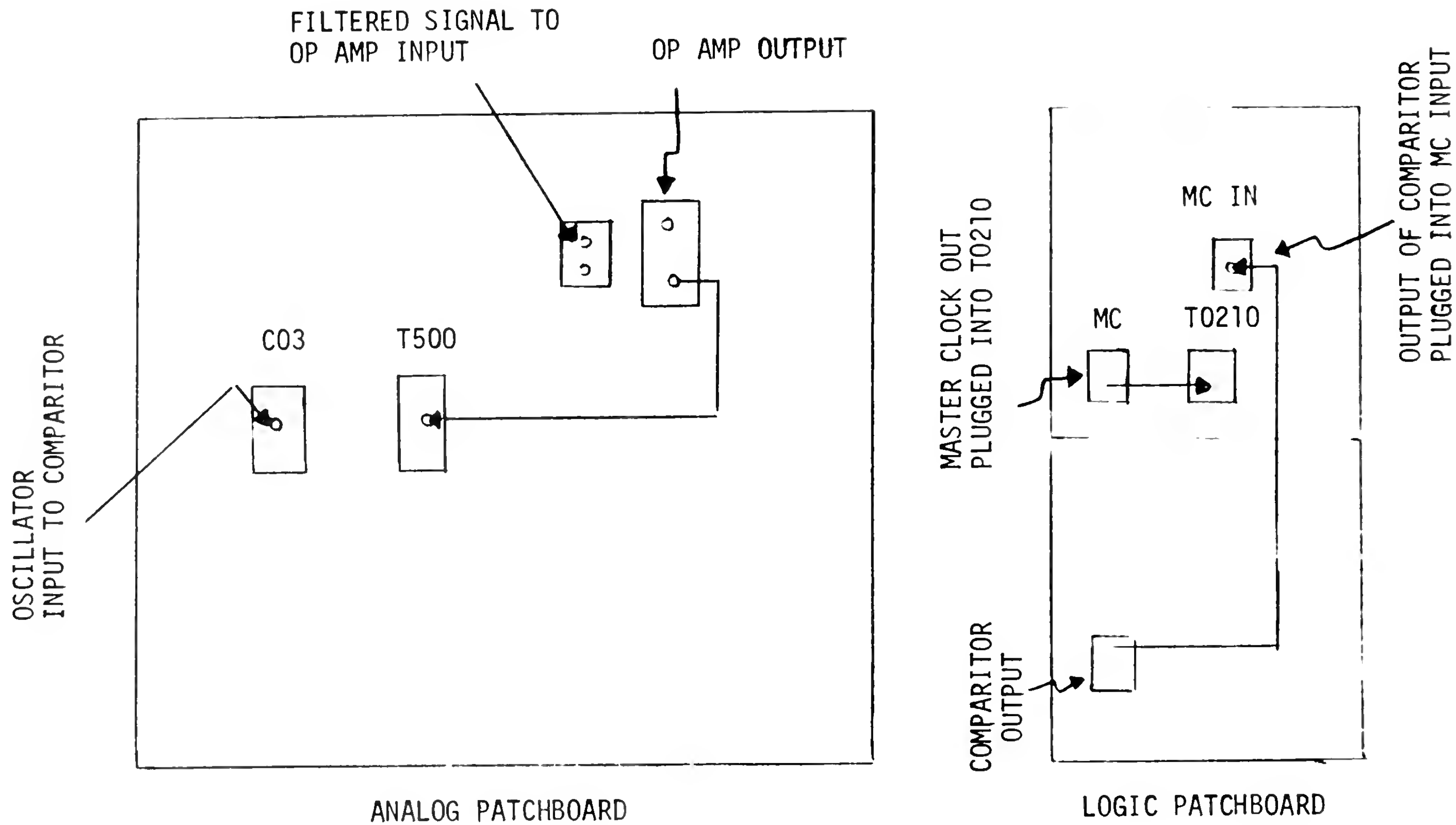
The program used to convert from seven track, 24 bits to nine track, 32 bit data is listed in Enclosure (3). The program is both Fortran IV and assembly language. The control cards are listed below.

Card No.

```
  1     //Name    (standard job card) ...

  2     //FOURIER EXEC FORTCALG,TIME.GO=05,REGION .GO=160K

  3     //FORT.SYSIN DD *

          .
          .       (main FORTRAN IV PROGRAM)
          .
        /*

  4     //ASM.SYSIN DD *

          .
          .       (ASSEMBLY LANGUAGE PROGRAM)
          .

  5     //GO.FT02F001 DD UNIT=0C0,VOL=SER=NPSYYY,LABEL=(,BLP),          *

  6     //               DISP=(OLD,KEEP),DCB=(DEN=0,RECFM=F,BLKSIZE=
        4096)

  7     //GO.FT04F001 DD UNIT=2400,VOL=SER=NPSXXX,LABEL=(,SL),DISP=
        (NEW,KEEP),      *

  8     //               DCB=(DEN=2,RECFM=F,BLKSIZE=4096),DSNAME=____
```

Card #1,2,3, are the standard control cards for the NPGS Computer facility. Card 4 indicates the following program is in assembly language. Card 5 and 6 tell the computer where the seven track data is located and its format. For example, the card above says the 7-track data is located on tape NPSYYY, to be placed on tape drive 02, the label processing is to be bypassed, with fixed format, with blocksize equal to 4096 bits, (1024 words). Cards 7 and 8 tell the computer where the data is to be transferred. The data is to be written on tape NPSXXX, tape drive 04, with standard label, fixed format in blocks of 4096 bits (1024 words). The 1024 words corresponds to the number of samples/record in the Subroutine SAMPL on the SDS 9300 when NSAMP=1024.

Anticipate difficulty in the conversion of the data. The tape drives are the weakest link and will require persistence to be successful. The first record of the data will probably be missed. The known data will allow a quick check of the system.



Enclosure(1) to Appendix B

Enclosure(2) to Appendix B

```
C
C          THIS PROGRAM SAMPLES
C        BOND AND MYATT    DIVERS SPEECH
C
C
C      * IN COLUMN 6 INDICATES CONTINUATION FROM PREVIOUS CARD
C       FOR THIS PROGRAM LIST ONLY
C
AFORTRAN GO,LS
      DIMENSION IBUF1(1024),IBUF2(1024),BUF(2048,2),JXY(2),Y
     *S(513)
      NAMELIST NCBLKS,LENGTH,ITAPE,ITOG
      ITAPE=1
      NOBLKS=2
      LENGTH=1024
      ITOG=0
      JXY(1)=1
      JXY(2)=2
 1    OUTPUT(101) NOBLKS,LENGTH,ITAPE,ITOG
      WRITE(102,104)
 104  FORMAT($ TYPE IN CHANGES.$/$ FOLLOW WITH AN * AND RETU
     *RN.$)
      INPUT(101)
C
 101  FORMAT(2I2)
 2    IF(ITAPE)3,4,3
 3    WRITE(102,102)ITAPE
 102  FORMAT($ CHECK ON TAPE$I1$ AND CLEAR PAUSE.$)
      PAUSE 10
 4    CALL SAMPL(NCBLKS,LENGTH,IBUF1,IBUF2,ITAPE,ITOG)
 5    IF(ITOG.EQ.0)GO TO 5
 6    CONTINUE
      NT=0
      IF(ITOG.EQ.-1)GO TO 10
 7    IF(ITOG.EQ.1)GO TO 7
 8    CONTINUE
      NT=1
      IF(ITOG.EQ.-1)GO TO 10
      GO TO 5
 10   WRITE(102,103)
 103  FORMAT(//5X$END RUN...$/$SET DESIRED SENSE SWITCH.$)
 20   PAUSE 20
      IF(SENSE SWITCH 3)11,12
 12   IF(SENSE SWITCH 4)13,14
 14   IF(SENSE SWITCH 5) 15,16
 16   IF(SENSE SWITCH 6)1,20
 11   ML=(1-NT)*LENGTH
      NL=NT*LENGTH
      DO 18 I=1,LENGTH
      BUF(I,2)=I
      BUF((I+LENGTH),2)=I+LENGTH
      BUF(I+ML,1)=IBUF1(I)
 18   BUF((I+NL),1)= IBUF2(I)
      L4=LENGTH/2
      L=LENGTH*2
      CALL LONGPLOT(BUF,JXY,L,2048,1,0,0.0,0.0,0.0,0.0,L4,YS
      GO TO 20
 13   ENDFILE ITAPE
      GO TO 20
 15   REWIND ITAPE
      GO TO 20
      END
AMETA9300   SI,LO,GO
$SAMPL PZE
*
*
*  THIS PROGRAM SAMPLES ONE ANALOG SIGNAL ON TRUNKLINE 0500
*  AT A RATE DETERMINED BY A CLOCK PULSE APPEARING ON TRUNKLI
     *NE
*  0210.  ANALOG TO DIGITAL CONVERSION IS PERFORMED AND THE
*  DATA IS STORED IN BUFFERS SET UP IN THE CALLING PROGRAM.
*  CALLING SEQUENCE IS...
```

```
*        CALL  SAMPL(NREC,NSAMP,IBURO,IBUF1,NTAPE,ITOG)
*  WHERE...
*       NREC=NUMBER OF RECORDS/RUN
*       NSAMP=NUMBER OF SAMPLES/RECORD
*       IBUFO=FIRST DATA BUFFER
*       IBUF1=SECOND DATA BUFFER
*       IBUFO=FIRST DATA BUFFER
*             BUFFERS APE FILLED WITH NSAMP A/D SAMPLES ALTERNA
      *TELY
*       NTAPE=MAG TAPE UNIT NUMBER
*             IF=0, NC TAPE IS WRITTEN
*       ITOG=0 IF IBUFO CURRENTLY BEING FILLED
*           =1 IF IBUF1 CURRENTLY BEING FILLED
*
*       AFTER INITIALIZING, THE SUBROUTINE WAITS FOR THE ANALO
      *G TO
*       GO TO COMPUTE.     END OF RUN IS CAUSED BY ANALOG GOIN
      *G OUT
*       OF COMPUTE OR WHEN NREC RECORDS OF DATA HAVE BEEN TAKE
      *N.
*       IN EITHER CASE, ITOG IS SET TO -1 TO NOTIFY THE CALLIN
      *G
*       ROUTINE OF THE ENDRUN CONDITION.
*       NOTE THAT AFTER INITIALIZING THE A/D SAMPLING, CONTROL
      *IS
*       RETURNED TO THE CALLING ROUTINE AND INTERRUPTS CALL TH
      *E
*       A/D INPUT ROUTINE.
*
ITG      OPD      0100000
         BRM      9SETUPN
         PZE      6
NREC     ITG                          NUMBER OF RECORDS/RUN
NSAMP    ITG                          NUMBER OF SAMPLES/RECORD
BUFFO    ITG                          FIRST DATA BUFFER
BUFF1    ITG                          SECOND DATA BUFFER
TAPE     ITG                          MAG TAPE UNIT NUMBER
TOGGL    ITG                          TOGGLE...INDICATES BUFFER BEIN
      *G FILLED
*                                         OR ENDRUN CONDITION.
A        EQU      5
NA       EQU      -5
B        EQU      4
X1       EQU      1
X2       EQU      2
X3       EQU      3
         EOM      033001              SETLINES 1 AND 2 FALSE
         POT      =0
         LDA      *TOGGL
         STA      TIME
         LDA      =-1
         STA      NOTAPE
         LDA      *TAPE
         ETR      =07
         SKU      =0
         STA      NOTAPE
         COPY     (A,B)
         MRG      WTBX
         STA      WTBN
         COPY     (B,A)
         MRG      EFTX
         STA      EFTN
         COPY     (B,A)
         MRG      TRTX
         STA      TRTN
         COPY     (B,A)
         MRG      BTTX
         STA      BTTN
         LDA      *NSAMP
         COPY     (A,B)
         SUB      =1
         MRG      PLACE
```

```
        STA     SDATAT
        STA     CURX2
        COPY    (0,A)
        LLSD    14
        STB     DAT
        MRG     ALC
        STA     ALC0
        STA     ALC1
        COPY    (0,A)
        LDB     BUFF0
        LLSB    9
        LLSD    1
        LRSB    10
        LLSA    5
        MRG     ALC0
        STA     ALC0
        COPY    (B,A)
        MRG     DAT
        STA     CW0
        COPY    (0,A)
        LDB     BUFF1
        LLSB    9
        LLSD    1
        LRSB    10
        LLSA    5
        MRG     ALC1
        STA     ALC1
        COPY    (B,A)
        MRG     DAT
        STA     CW1
        LDA     BUFF0
        ETR     =077777
        ADD     =-1
        STA     ORIG0
        STA     COMM
        LDA     BUFF1
        ETR     =077777
        ADD     =-1
        STA     ORIG1
        LDA     OCW
        MRG     =COMM
        STA     SVCW
        STA     CW
        LDA     BRMAD
        XMA     040
        STA        SV040
        LDA     BRMSAM
        XMA     051
        STA     SV051
        SKN     NOTAPE
        BRU     AGAIN
        LDA     BRMFL
        XMA     010
        STA     SV010
        LDA     BRMFL
        XMA     011
        STA     SV011
        EXU     TRTN
        CAT     0
        BRU     $-2
        EXU     BTTN
        BRU     $+2
        BRU     AGAIN
        EOM     *014000
        POT     SPACE
        EXU     EFTN
AGAIN   STZ     *TOGGL
        STZ     TOGGT
        LDA     =-1
        STA     ADEL
        LDA     *NREC
        ADD     =-1
```

```
        STA     COUNT
        EOM     033001
        POT     =040000000       SETLINE 1 TRUE
        SKS     030010  ANALOG IN COMPUTE TEST
        BRU     $-1
        EOM     033001           FALSE
        POT     =0
        EIR
        EOM     033004           ENABLE PATCHBOARD INTERRUPTS
        BPR     SAMPL            RETURN TO MAIN PROGRAM
* END OF INITIAL SET UP
* SUBROUTINE START ENTERED ON INTERRUPT 051
*
START   PZE
        STA     SVA
        STX     SVX2,X2
        SKN     ADFL
        BRU     $-1
        DIR
        LDA     *TOGGL   THIS PROCEDURE TESTS
        SKG     TOGGT   TO DETERMINE WHICH
        BRU     LOAD0   BUFFER TO LOAD
LOAD1   MPO     COMM
        SKS     030010           ANALOG IN COMPUTE
        BRM     ENDRUN             NO
        STZ     ADFL
        EOM     034000
        POT     CW
        EIR
        LDX     CURX2,X2
RTEST1  BRX     RIDLE,X2
        SKN     ADFL
        BRU     $-1
        LDA     ORIGO
        STA     COMM
        LDA     SVCW
        STA     CW
        STZ     *TOGGL
        LDX     SDATAT,X2
RTEST2  SKN     NOTAPE
        BRU     INCR
        EXU     TRTN
        CAT     0
        BRU     $-2
        LDA     TIME
        ADM     *BUFF1
        ADD     =1
        SKL     =1000
        COPY    (0,5)
        STA     TIME
        EXU     WTBN
        EXU     ALC1
        POT     CW1
        BRU     INCR
*
*
*  THIS DIVIDES THE SUBROUTINE INTO BUFFERS
*
LOAD0   MPO     COMM
        SKS     030010           ANALOG IN COMPUTE
        BRM     ENDRUN             NO
        STZ     ADFL
        EOM     034000
        POT     CW
        EIR
        LDX     CURX2,X2
RTEST3  BRX     RIDLE,X2
        SKN     ADFL
        BRU     $-1
        LDA     ORIG1
        STA     COMM
        LDA     SVCW
```

```
         STA      CW
         MRG      *TOGGL
         LDX      SDATAT,X2
RTEST4   SKN      NOTAPE
         BRU      INCR
         EXU      TRTN
         CAT      0
         BRU      $-2
         LDA      TIME
         ADM      *BUFFO
         ADD      =1
         SKL      =1000
         COPY     (0,5)
         STA      TIME
         EXU      WTRN
         EXU      ALCO
         POT      CWO
INCR     SKR      COUNT     BLOCK COUNT IS REDUCED.
         BRU      RIDLE     IF NEGATIVE, ALL DATA HAS BEEN
*C                          TRANSFERRED.
         BRU      FIN
RIDLE    LDA      SVA
         STX      CURX2,X2
         LDX      SVX2,X2
         BRC      *START
FIN      EOM      033000
         DIR
         EOM      033001              SETLINE 2 TRUE
         POT      =020000000
         SKS      031001              WAIT FOR ANALOG
         BRU      $-1                 ON TEST 1.
         EOM      033001              SETLINE 2 FALSE
         POT      =0
         SKR      *TOGGL
         BRU      $-1
         CAT      0
         BRU      $-1
         LDA      SV040
         STA      040
         LDA      SV051
         STA      051
         LDA      SV010
         STA      010
         LDA      SV011
         STA      011
         EIR
         BRU      RIDLE
ENDRUN   PZE
         DIR
         LDA      COMM
         ETR      =077777
         STA      COMM
         LDA      =-1
         LDX      CURX2,X2
         STA      *COMM
         MRG      COMM
         BRX      $-2,X2
         STZ      COUNT
         LDA      ENDRUN
         SKG      =LOADC
         BRU      RTEST2
         BRU      RTEST4
*
*    END OF  SUBROUTINE
BRMPL    BRM      PLUG
PLUG     PZE
         BRC      *PLUG
SV010    PZE
SV011    PZE
*
CON      FORM     9,15
CONT     FORM     10,14
```

```
SPACE   CONT    150,0
OCW     DATA    0100000
SVCW    PZE
CW      PZE
COMM    PZE
        PZE
SDATAT  RES     1
ORIG0   RES     1
ORIG1   RES     1
DAT     RES     1
ALC     EOM     014000
ALC0    PZE
ALC1    PZE
CW0     RES     1
CW1     RES     1
TOGGT   RES     1
COUNT   RES     1
SVA     PZE
SVX2    PZE
CURX2   PZE
SV040   RES     1
SV051   RES     1
PLACE   DATA    077700000
TIME    PZE
NOTAPE  PZE
BTTX    BTT     0,0
BTTN    PZE
WTBX    WTB     *0,0,4
WTBN    PZE
EFTX    EFT     0,0,4
EFTN    PZE
TRTX    TRT     0,0
TRTN    PZE
BRMSAM  BRM     START
BRMAD   BRM     ADEND
ADEND   PZE
        SKR     ADFL
        BRU     $-1
        BRC     *ADEND
ADFL    PZE
        END
4EOF
4LOAD XM,MAP.
Δ        DATA
```

Enclosure(3) To Appendix B

```
          THIS PROGRAM INTERFACES THE SDS 9300 TO THE IBM360



//MYA00053 JOB (0053,47FP),'MYATT      ',MSGLEVEL=1,CLASS=G
//FOURIER EXEC FRTCALGP,TIME.GO=10,REGION.GO=160K
//FORT.SYSIN DD *
      DIMENSION INDATA(1024),DATA(1024)
      FACTOR=100.0/(2**31-1)
      REWIND 2
      REWIND 4
     DO J=1,K     WHERE K IS THE NUMBER OF BLOCKS OF 1024 FAC
      DO 31 J=1,133
      READ(2,3,END=40,ERR=50) INDATA
    3 FORMAT(16(64A4))
      CALL FORM(INDATA)
      WRITE(6,70) J
   70 FORMAT ('1',10X,'RECORD NO.=',I6)
      DO 1 I=1,1024
    1 DATA(I)=INDATA(I)*FACTOR
      WRITE(6,66) (DATA(I),I=1,1024)
   66 FORMAT(1X,10E10.2)
      WRITE (4,3) DATA
      GO TO 31
   50 WRITE (6,51) J
   51 FORMAT ('0',5X,'READ ERROR, RECORD NO.=',I6)
   31 CONTINUE
      STOP
   40 WRITE (6,41) J
   41 FORMAT ('0',5X,'END OF TAPE, RECORD NO.=',I6)
      STOP
      END

//ASM.SYSIN  DD   *
FORM       START 0
*          SUBROUTINE FORM(INDATA)
*
*     * INDATA OF AN ARRAY LENGTH SPECIFIED BY THE INDEX VALU
* THIS SUBROUTINE WILL CONVERT 24 BIT BINARY WORDS STORED IN
*     $E TO 32 BIT BINARY WORDS AND PLACE THESE SAME WORDS BA
*     $CK INTO INDATA
*
*
*
           STM    14,12,12(13)        THIS SUBROUTINE CONVERTS
           BALR   6,0                 24 BIT BINARY WORDS TO
           USING  *,6                 32 BIT WORDS
           USING  DATA,7
           SR     7,7
           L      11,=F'1024'       THIS IS THE INDEX VALUE
           L      12,0(1)
 LOOP      L      2,NUM(12)
           LR     3,7
           SRDL   2,6
           SRL    2,2
           SRDL   2,6
           SRL    2,2
           SRDL   2,6
           SRL    2,2
           SRDL   2,6
           ST     3,NUM(12)
           LA     12,4(12)
           BCT    11,LOOP
           LM     2,12,28(13)
           MVI    12(13),X'FF'
           BCR    15,14
DATA       DSECT
NUM        DS     1F
           END

//GO.FT02F001 DD UNIT=OCO,VOL=SER=DIVERS,LABEL=(,BLP),
    $                 *
```

```
//                DISP=(OLD,KEEP),DCB=(DEN=0,RECFM=F,BLKSIZE=40
     $96)
//GO.FT04F001 DD UNIT=2400,VOL=SER=NPS234,LABEL=(,SL),DISP=(
     $NEW,KEEP),  *
//                DCB=(DEN=2,RECFM=F,BLKSIZE=4096),DSNAME=MYA00
     $053
```

COMPUTER PROGRAM HARM

PROGRAM COMPUTES POWER SPECTRUM FOR 8192 DATA POINTS

```
C
C
C     $ IN COLUMN 6 INDICATES CONTINUATION FROM PREVIOUS CARD
C     FOR THIS PROGRAM LIST ONLY
C
//MYA30053 JOB (0053,47FP),'MYATT     ',MSGLEVEL=1,CLASS=D
//FOURR EXEC FORTCLGP,PARM.FORT='LINECNT=75',TIME.GO=02,REGI
     $ON.GO=300K
//FORT.SYSIN DD *
      REAL*4LABEL(10)/40HONE TWO THREFOURFIVE SIXSEVNEIGTNIN
     $E TEN     /
      REAL*8ITITLE(12)/96H      POWER SPECTRUM      J.M. MYATT
     1    M**2  111-118 W   MUZZLE SINGING
      DIMENSION XG(420),D(1260),Y(8200),B(1260)
      DIMENSION S(8200),INV(8200),M(3),DATA(1024)
      COMPLEX*8 A(8192,1,1),Z(1260,1,1)
      CALL CANCEL(2)
C
C
      ITYPE=2
      KK=13
      TIME =ITIME(0)*0.01
      N1=2**KK
      NPT=N1
      NY=1
      NN=420
      NZT=420
C                     PROGRAM COMPUTES 8192 COEF.
C      PLOT GOES TO 3000 HERTZ
      IMAX=1
      BMAX=0.0
      KA=2
      KC=2
      KO=2
C      SAMPLING RATE IS 2048 CPS.
      XN=N1
      DT=1.0/2048.0
      BW= 1.0/(2.0*DT)
      DELTF= 1.0/(XN*DT)
      N2=1
      N3=1
      M(1)=KK
      M(2)=0
      M(3)=0
      KB=0
C                     NN= RUNNING COUNT OF POINTS BEING PLOTTED
C                     NZT= NUMBER OF POINTS PER GRAPH
C                     DELTF = FREQUENCY RESOLUTION
C
C
C        J IS # OF BLOCK OF DATA ON MAG TAPE BEING READ
C
      DO 700 J=1,118
      IO=KB*1024
      READ (4,50) (DATA(I),I=1,1024)
   50 FORMAT(16(64A4))
      IF(J.LT.111)GO TO 700
      DO 777 I=1,1024
  777 Y(IO+I)=DATA(I)
      KB=KB+1
  700 CONTINUE
      DO 2 I1=1,N1
      DO 2 I2=1,N2
      DO 2 I3=1,N3
    2 A(I1,I2,I3)=Y(I1)
C
      WRITE(6,40)DT,J,N1,NZT
   40 FORMAT(////3X,'SAMPLING INTERVAL= ',F10.6,10X,' SOUND
     $/AE/      '
     *,10X,' BLOCK = ',I6,5X,'DATA POINTS',2I7/)
      WRITE(6,66)BW,DELTF
```

```
   66 FORMAT(//6X,'BANDWIDTH = ',F15.6,10X,'DELTA F = ',F10.
     $6)
      WRITE(6,10)(((I1,      A(I1,I2,I3),I3=1,N3),I2=1,N2),I
     $1=1,N1)
   10 FORMAT(//' AMPLITUDE OF FOURIER FUNCTION TO BE TRANSFO
     $RMED'/(4(1X,
     *'A(',I5,      ')=',1X,2F7.3,4X)))
      CLOCK=ITIME(0)*0.01
      CALL HARM(A,M,INV,S,-1,IFERR)
      DO 34 I1=1,NZT
   34 XG(I1)=I1-1
C   KO IS TEST FOR SMOOTHING
      IF(KO.LT.1)GO TO 800
      Z(1,1,1)=A(1,1,1)
      B(1)=CABS(Z(1,1,1))
      DO 801 I1=2,1260
      I2=I1-1
      I3=I1+1
      Z(I1,1,1)=-0.25*(A(I2,1,1)+A(I3,1,1))+0.5*A(I1,1,1)
      B(I1)=CABS(Z(I1,1,1))
      B(I1)=B(I1)**2
  801 CONTINUE
      WRITE(6,965)
  965 FORMAT(1H0,'  FIRST SMOOTHING FUNCTION USED ')
      GO TO 820
  800 DO 33 I1=1,1260
      B(I1)=CABS(A(I1,1,1))
   33 B(I1)=B(I1)**2
  820 DO 11 I1=1,420
   11 XG(I1)=XG(I1)*DELTF
      J1= 0
C     KC IS SECOND SMOOTHING FUNCTION CHECK
C
      IF(KC.GT.1)GO TO 987
  990 DO 666 I1=1,1260
      IF(B(I1).LT.BMAX)GO TO 666
      BMAX=B(I1)
      IMAX=I1
  666 CONTINUE
      WRITE(6,667)BMAX,IMAX
  667 FORMAT(1H0,'    BMAX = ',F7.3,'  IMAX = ',I6)
      DO 78 J3=1,3
      J2= 1
C    KA IS CHECK FOR LOG10 PLOT
      IF(KA.LT.1)GO TO 840
      DO 850 I1=NY,NN
      B(I1)=B(I1)/BMAX
      IF(B(I1).LE.0.002)GO TO 1
      D(J2)=10.0*ALOG10(B(I1))
      IF(D(J2).GE.-27.0)GO TO 850
    1 D(J2)=-27.0
  850 J2=J2+1
  851 CALL DRAW(NZT,D,XG,0,0,LABEL(J3),ITITLE,3.,7.0,0,0,0,0
     $,9,15,1,LAST)
      GO TO 855
  840 DO 77 I1=NY,NN
      D(J2)=B(I1)
      J2=J2+1
   77 CONTINUE
      CALL DRAW(NZT,D,XG,0,0,LABEL(J3),ITITLE,0,7.0,0,0,0,0,
     $9,15,1,LAST)
C     NOTE$$$$$$ X AND Y AXIS ARE REVERSED
C
  855 WRITE(6,20)(((I1,      A(I1,I2,I3),I3=1,N3),I2=1,N2),I
     $1=NY,NN)
   20 FORMAT(//'FOURIER TRANSFORM A(I1,I2,I3)'/(4(1X,'A('I5,
     $     ')=',1X
     *,  2F7.3,4X)))
      NY=NN+1
      NN=NN+420
   78 CONTINUE
      ITYPE=2
```

```
      GO TO 35
  987 D(1)=B(1)
      DO 988 I1=2,1259
      I2=I1-1
      I3=I1+1
      D(I1)= 0.5*B(I1)+0.25*(B(I2)+B(I3))
  988 CONTINUE
      DO 989 I1=1,1259
      B(I1)=D(I1)
  989 CONTINUE
      WRITE(6,966)
  966 FORMAT(1H0,'    SECOND SMOOTHING FUNCTION USED')
      GO TO 990
   35 CONTINUE
      END
//GO.FT04F001 DD DSNAME=MYA0053,UNIT=2400,DCB=(RECFM=F,BLKSI
      SZE=4096),    *
//                  VOLUME=(PRIVATE,SER=NPS234)
```

BIBLIOGRAPHY

1. Communication Sciences Laboratory/Office of Naval Research Report 20, Abstracts of Four Studies on Underwater Communications, by H. Hollien, J. Brandt and J. Malone, 1967-68.

2. Special Issue on Fast Fourier Transform, - Institute of Electronics Engineers Transactions on Audio and Electroacoustics, June 1967.

3. Hydro-Products Division of Dillingham Company, Facts on Underwater Communications, 1969.

4. CSL/ONR Report 15, The Effect of Air Bubbles in The External Auditory on Underwater Hearing Thresholds, by H. Hollien, T.E. Doherty and J. Brandt, March 1967.

5. CSL/ONR Report 11, Intelligibility of Diver Communication Systems, by H. Hollien, J. Brandt and T.E. Doherty, 15 December 1968.

6. Fant, C.G.M., "Speech at High Ambient Pressures," Acoustic Specification of Speech, Royal Institute of Technology, Speech Transmission Laboratory, Sweden.

7. Mullen, W.W.Jr., "Literature Survey Pertinent to Underwater Voice Communication," Technical Notes (Unpublished), U.S. Navy Mine Defense Laboratory, Panama City, Florida.

8. Brubaker, R.S. and Wurst, J.W., "Spectrographic Analysis of Divers' Speech During Decompression," Journal of Acoustical Society of America, v. 43, no. 4, 1968.

9. Potter, R.K. and Kopp, G.A., Visible Speech, Dover Publications, Inc., New York, 1966.

10. Comer, D.J., "The Use of Waveform Asymmetry To Identify Voiced Sounds," IEEE Transactions on Audio and Electroacoustics, v. AU-16, no. 4, December 1968.

11. Lindgren, Nilo, "Speech-Mans Natural Communication," IEEE Spectrum, v. 4, no. 6, June 1967.

12. Oppenheim, A.V., "Speech Analysis-Synthesis System Based on Homomorphic Filtering," JASA, v. 45, p. 458-465, February 1969.

13. Priestly, M.B., "Power Spectral Analysis of Non-Stationary Random Processes," Journal of Sound and Vibrations, v. 6, p. 86-97, July 1967.

14. Dorrian, L.V., "Digital Spectral Analysis," Masters Thesis, Naval Postgraduate School, Monterey, California, September 1968.

15. Post, J.L., "Analysis and Synthesis of a Time Limited Complex Wave Form," Masters Thesis, Naval Postgraduate School, Monterey, California, December 1968.

16. Bingham, C., Godfrey, M.D., and Tukey, J.W., "Modern Techniques of Power Spectrum Estimation," IEEE Transactions on Audio and Electroacoustics, v. AU-15, no. 2, June 1967.

17. Singleton, R.C. and Poulter, T.C., "Spectral Analysis of the Call of the Male Killer Whale," IEEE Transactions on Audio and Electroacoustics, v. AU-15, no. 2, June 1967.

18. Bremerman, H.J., "Pattern Recognition Functional and Entropy," Contract NONR 3656 (08) and NONR 222 (88), December 10, 1967.

19. Gold, B. and Rabiner, L.R., "Analysis of Digital and Analog Format Synthesizer," IEEE Transactions on Audio and Electroacoustics, v. AU-16, p. 81-94, March 1968.

20. Fischer-Jorgenson, E., "What Can the New Techniques of Acoustic Phonetics Contribute to Linguistics," Psycholinguistics, Holt, Rinehart and Winston, New York, 1961.

21. Morrow, C.T., "Reaction of Small Enclosures on the Human Voice, Part II," JASA, v. 20, p. 487-97, July, 19~~6~~48.

22. Drucher, H., "Speech Processing in High Ambient Noise Environment," IEEE Transactions on Audio and Electroacoustics, v. AU-16, no. 2, p. 166, June 1968.

23. Hunter, E.K., "Problems of Diver Communication," IEEE Transactions on Audio and Electroacoustics, v. AU-16, no. 1, p. 118-21, March 1968.

24. Kenny, J.E., "Some Design Problems in Wireless Diver Communications," IEEE Transactions on Audio and Electroacoustics, v. AU-14, no. 4, p. 174-177, December 1966.

25. Morrow, C.T., "Reaction of Small Enclosures on the Human Voice, Part I," JASA, v. 19, p. 645-652, July 1947.

26. Olson, H.F., and others, "Speech Processing Techniques and Applications," IEEE Transactions on Audio and Electroacoustics, v. AU-15, no. 3, September 1967.

27. Parzen, E., "Informal Comments on Uses of Power Spectrum Analysis," IEEE Transactions on Audio and Electroacoustics, v. AU-15, no. 2, p. 75-76, June 1967.

28. Pickett, J.M., "Low Frequency Noise and Methods for Calculating Speech Intelligibility," JASA, v. 31, September 1959.

INITIAL DISTRIBUTION LIST

		No Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	20
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Professor Harold A. Titus, Code 52TS Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
4.	Maj. William A. Bond, USMC Box 116 Nocona, Texas 76255	1
5.	Capt. J.M. Myatt, USMC 2027 Ferndale Avenue Dallas, Texas 75224	1
6.	Commandant of the Marine Corps(Code A03C) Headquarters, U. S. Marine Corps Washington, D. C. 20380	1
7.	James Carson Breckinridge Library Marine Corps Development and Educational Command Quantico, Virginia 22134	1
8.	Dr. Harry Hollien Communications Sciences Laboratory Department of Speech University of Florida Gainesville, Florida 32601	1
9.	Mr. James Lasch Hydro Products Divison of Dillingham Corporation Box 2528 San Diego, California 92112	1
10.	Mr. Omar Lamborn Electro-Dynamics Divison Bendix Corporation 11600 Sherman Way North Hollywood, California 92112	1

No. Copies

11. LCDR Al Spinks 1
Naval Special Warfare Group
Naval Amphibious Base
Coronado, California

12. Mr. Monns Turntine 1
Naval Applied Science Laboratory
Flushing and Washington Avenues
Brooklyn, New York 11251

Security Classification

DOCUMENT CONTROL DATA - R & D			
(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)			
1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940		2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP	
3. REPORT TITLE Investigation of Distortion of Divers' Speech Using Power Spectral Estimates Based on the Fast Fourier Transform			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Master's thesis, June 1969			
5. AUTHOR(S) (First name, middle initial, last name) William H. Bond James M. Myatt			
6. REPORT DATE June 1969	7a. TOTAL NO. OF PAGES 96		7b. NO. OF REFS 28
8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)		
10. DISTRIBUTION STATEMENT This document has been approved for public release and sale; its distribution is unlimited.			
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Postgraduate School Monterey, California 93940		

13. ABSTRACT

The problem of distortion in underwater communications peculiar to free divers and techniques for analysis of speech wave forms are discussed. The Fast Fourier Transform algorithm, selected to analyze shifts in formant frequencies due to restricted oral cavities, high ambient pressures, and forced speech is discussed. The Fast Fourier Transform is used to analyze a vowel sound and show that the expected shifts do occur. Recommendations are made for extending the techniques to all non-noise like sounds and breathing mixtures other than compressed air.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Security Classification

A-31408

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Underwater communication						
Divers' speech						
Fast Fourier Transform						
Power spectral estimates						
Speech processing						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0101-807-6821